



Nuria Oliver

## Inteligencia Artificial: ficción, realidad y... sueños



Nuria Oliver es Directora de Investigación en Ciencias de Datos en Vodafone. Es ingeniera superior de Telecomunicación por la UPM y doctora por el Massachusetts Institute of Technology en Inteligencia Perceptual. Tiene más de 20 años de experiencia investigadora en MIT, Microsoft Research y como primera directora científica en Telefónica I+D y en Vodafone.

Nuria asesora al Gobierno de España y a la Comisión Europea en Inteligencia Artificial, a Mahindra Comviva y al Future Digital Society. Además, es la única investigadora española reconocida por el ACM como Distinguished Scientist, Fellow y miembro de CHI Academy. También es Fellow del IEEE y de la Asociación Europea de Inteligencia Artificial. Doctora Honoris Causa por la UMH y Académica de la Real Academia de Ingeniería y de la Academia Europea.

Ha sido nombrada una de las 11 personas más influyentes de Inteligencia Artificial en el mundo por *Pioneering Minds* (2017), una de las mentes maravillosas españolas en tecnología (EL PAIS, 2017), "una destacada directora de tecnología" (EL PAIS, 2012), uno de los "100 líderes para el futuro" (Capital, 2009) y uno de los "40 jóvenes que marcarán el próximo milenio" (EL PAIS, 1999), entre otros.



Reproducción autorizada de la obra  
"Inteligencia Artificial: Ficción, Realidad y ...Sueños"

Editada por la Real Academia de Ingeniería  
© 2018, Real Academia de Ingeniería  
© 2018 del texto, Nuria María Oliver Ramírez  
ISBN: 978-84-95662-64-4  
Depósito legal: M-38898-2018  
Impreso en España



Nuria Oliver

# Inteligencia Artificial: ficción, realidad y... sueños

Discurso de toma de posesión como  
académica de número de la  
**Excma. Sra. D.A Nuria M. Oliver Ramírez**

Leído en la sesión solemne de su toma  
de posesión que tuvo lugar  
el día 11 de diciembre de 2018

Editado e impreso por primera vez por la  
**Real Academia de Ingeniería**  
Madrid



Presentación de Fernando Plaza, Director de gtt	9
1. Breve Historia de la Inteligencia Artificial (IA): Mi Visión Personal	31
2. Tipos de Inteligencia Artificial	33
3. Inteligencia Artificial versus Inteligencia Humana	37
4. La cuarta revolución industrial	
1. Transversalidad	
2. Complejidad, escalabilidad y actualización constante	
3. Habilidad para predecir	
5. El impacto laboral y económico de la Inteligencia Artificial	43
6. El potencial de la Inteligencia Artificial para mejorar la sociedad	49
7. Principios éticos y dimensiones de trabajo necesarias para el desarrollo de una Inteligencia Artificial centrada en las personas	59
8. Tres sueños para el futuro	73
9. Introducción al discurso de recepción de condición de Académica en la Real Academia de Ingeniería	75
10. Bibliografía	81
11. Curriculum Vitae	85



**Fernando Plaza**  
Director General de gtt

En gtt trabajamos intensamente por hacer realidad la Transformación Digital de las Administraciones Públicas. Desde nuestros inicios, hace más de dos décadas, hemos trabajado con una filosofía y unos principios que persiguen la eficacia y la eficiencia en la gestión de tributos de las corporaciones locales. Un escenario que nos lleva a recibir y valorar aportaciones de expertos que consideramos fundamentales para la revolución tecnológica a la que gratamente asistimos y en la que participamos

Estar en contacto con la innovación y la mejora continua de procesos nos hace ser conscientes del momento de cambio que vivimos. Lo que muchos llaman cuarta revolución industrial anticipa una nueva forma de vivir, trabajar y relacionarnos que lleva el concepto de avance tecnológico a un cambio de paradigma del que podrán beneficiarse los que sean capaces de innovar y adaptarse.

En este contexto, es muy grato para mí, poder editar y hacer extensivo el trabajo: "Inteligencia Artificial: ficción, realidad y... sueños" de la investigadora alicantina Nuria Oliver. Un trabajo destacable tanto por la calidad y la inteligencia del análisis, reflexión y planteamiento del contenido, como por ser éste el Discurso de su toma de Posesión como Académica Electa en la Real Academia de Ingeniería, el pasado 11 de diciembre de 2018. Un momento solemne y emblemático en el que tuve el honor de contarme entre los asistentes.

Nuria Oliver es una de las investigadoras más prestigiosas a nivel mundial. Es conocida por su trabajo en modelos computacionales de comportamiento humano, inteligencia artificial, interacción persona-máquina, informática móvil y Big data para bien social. Su extensa trayectoria tanto académica como en el terreno práctico y empresarial, la convierte en un referente de conocimiento clave para entender la nueva realidad. Por ello, recomiendo la experiencia de leer esta amena y clarificadora visión de la Inteligencia Artificial, que incluye aspectos históricos, tecnológicos y fundamentalmente, éticos. Una lectura interesante que abre la puerta a la reflexión acerca de las consecuencias de deconstruir la realidad tal y como la entendemos.

# 1. Breve Historia de la Inteligencia Artificial (IA): Mi Visión Personal

Aunque la Inteligencia Artificial (IA) parezca algo novedoso, el desarrollo de máquinas capaces de pensar o dotadas de algunas de las capacidades humanas ha cautivado nuestro interés desde la antigüedad. Los primeros autómatas —robots antropomorfos— que imitaban movimientos humanos fueron construidos hace milenios. Según la Ilíada, Hefesto —el dios griego del fuego y la forja— creó dos mujeres artificiales de oro con “sentido en sus entrañas, fuerza y voz” que lo liberaban de parte de su trabajo, es decir, robots para ayudarlo, lo cual lo convierte en todo un adelantado a su tiempo. Otros ejemplos incluyen autómatas con fines religiosos —por ejemplo, las figuras mecánicas de los dioses en el Antiguo Egipto para sorprender a la multitud con gestos operados por sacerdotes— y lúdicos —como las famosas “cabezas parlantes” y autómatas de la Edad Media, Renacimiento y el siglo XVIII.

Mas allá de la automatización, el ser humano siempre ha sentido curiosidad por explicar y entender cómo funciona la mente humana, entre otros motivos, como requisito necesario para poder construir una mente artificial. Hace más de 700 años, Ramon Llull —beato y filósofo mallorquín patrón de los informáticos— describió en su Ars Magna (1315) la creación del Ars Generalis Ultima, un artefacto mecánico capaz de analizar y validar o invalidar teorías utilizando la lógica; es decir, describe un sistema de Inteligencia Artificial.

Ada Byron (Lovelace) fue la primera persona en proponer en el siglo XIX el uso de la máquina analítica de Babbage para resolver problemas complejos más allá de cálculos matemáticos. Sin embargo, el mito y la ficción literaria respecto a la Inteligencia Artificial comenzaron a convertirse en realidad a partir de los años 40 del siglo XX con el desarrollo de los primeros ordenadores.

**Se considera que el genial matemático e informático inglés Alan Turing es el padre de la Inteligencia Artificial**, de la que habló en el legendario artículo “**Computing Machinery and Intelligence**”, publicado en 1950. En dicho artículo propuso **la famosa prueba de Turing**, ilustrada en la Figura 1, considerada el punto de referencia para determinar si un sistema artificial es inteligente. Esta prueba consiste en que un humano (C en la figura) y conocido como el “interrogador” interaccione vía texto con un sistema al que puede hacerle preguntas. El sistema se considera que supera la prueba de Turing si para el interrogador es imposible discernir cuándo las respuestas a sus preguntas son contestadas por una máquina (A en la figura) y cuándo lo son por un humano (B en la figura).

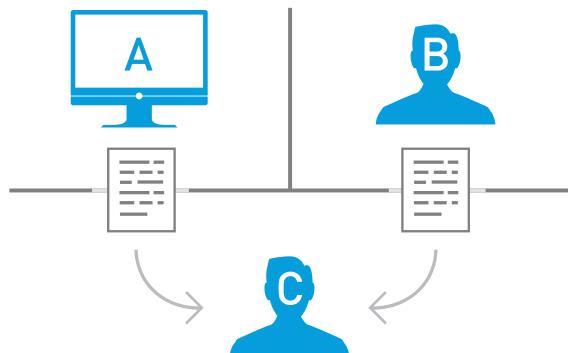


FIG 1. ESQUEMA DE LA PRUEBA DE TURING (FUENTE: WIKIPEDIA)

## La década de los 50 fue una década fundacional para la Inteligencia Artificial.

En 1951, el profesor Marvin Minsky —a quien tuve el honor de conocer en MIT— construyó la primera red neuronal computacional como parte de su doctorado en Princeton. En 1956 tuvo lugar la mítica convención de Dartmouth en la que participaron figuras legendarias como John McCarthy, Marvin Minsky, Claude Shannon, Herbert Simon y Allen Newell, y donde se definió el término de Inteligencia Artificial y se establecieron las bases para su desarrollo.

En este congreso, McCarthy definió la Inteligencia Artificial como “la disciplina dentro de la Informática o la Ingeniería que se ocupa del diseño de sistemas inteligentes”, esto es, sistemas con la capacidad de realizar funciones asociadas a la inteligencia humana, como aprender, entender, adaptarse, razonar e interactuar imitando un comportamiento humano inteligente. McCarthy acuñó el término para diferenciarlo del concepto de cibernetica, impulsado por Norbert Wiener (profesor de MIT) y con una visión de sistemas inteligentes basados en el reconocimiento de patrones, la estadística, las teorías de control y de la información. McCarthy quería enfatizar la conexión de la Inteligencia Artificial con la lógica. En un guiño del destino, la propuesta intelectual de Wiener se ha convertido en la aproximación dominante a la Inteligencia Artificial, pero utilizando la terminología de McCarthy. Sin embargo, no adelantemos acontecimientos...



FIG 2. PARTICIPANTES DE LA CONVENCIÓN DE DARTMOUTH, INCLUIDOS MARVIN MINSKY, CLAUDE SHANNON Y RAY SOLOMONOFF (FUENTE: CORTESÍA DE LA FAMILIA MINSKY)

En 1957 Rosenblatt propuso el Perceptrón en el Laboratorio Aeronáutico de Cornell. Una de sus primeras implementaciones se llamaba el Perceptrón Mark 1 (Figura 4), cuya aplicación era el análisis de imágenes. Rosenblatt, que era psicólogo, estaba convencido que el Perceptrón era una versión simplificada de cómo funcionan las neuronas. Como puede observarse en la Figura 3, el Perceptrón recibe un conjunto de valores de entrada (activaciones de neuronas adyacentes), multiplica cada valor de entrada por un peso (que representa la fuerza de la sinapsis con cada neurona adyacente), y produce una salida que es un 1 si la suma de las entradas moduladas por sus pesos es superior a un cierto valor, y 0 si es inferior (representando por tanto si esa neurona se activa o no). Este modelo está basado en trabajo anterior de McCulloch y Pitts que demostraron que un modelo de neurona como el descrito anteriormente puede representar funciones de OR/AND/NOT. Este resultado era importante porque en los albores de la Inteligencia Artificial se pensaba que cuando los ordenadores fuesen capaces de llevar a cabo operaciones de razonamiento lógico formal, se conseguiría alcanzar la Inteligencia Artificial.

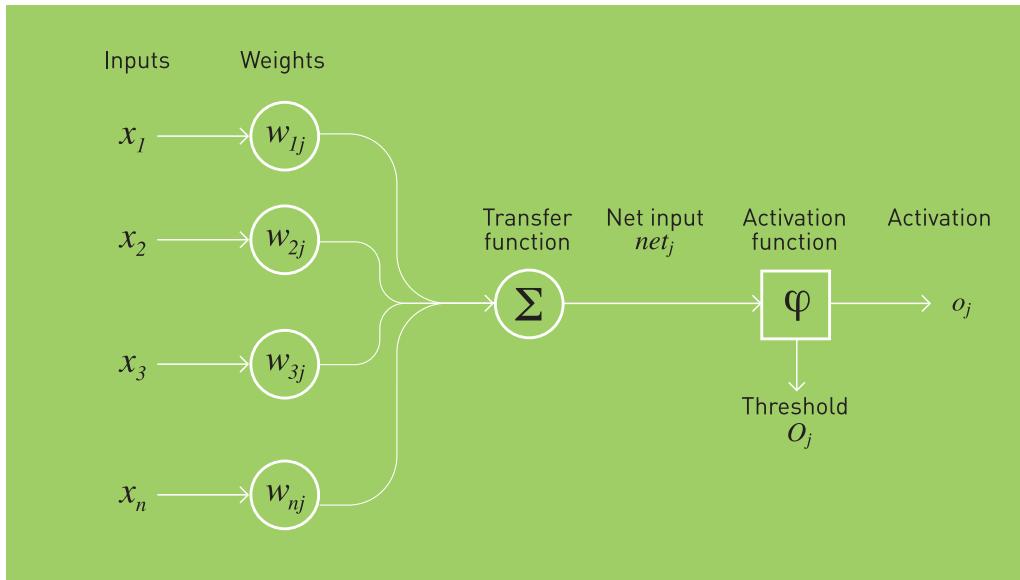


FIG. 3. PERCEPTRÓN DE ROSENBLATT (FUENTE: WIKIPEDIA)

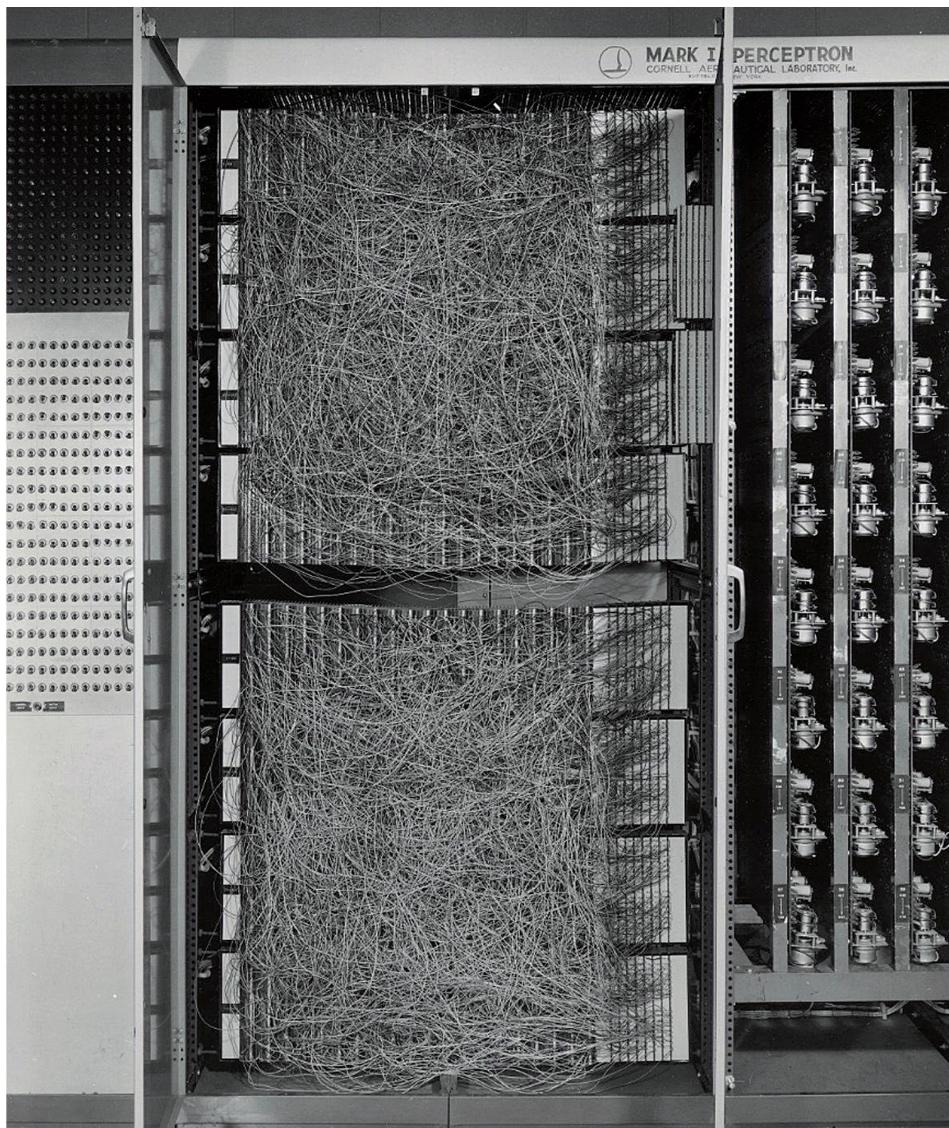


FIG. 4. PERCEPTRÓN MARK 1 DEL LABORATORIO AERONÁUTICO DE CORNELL  
(FUENTE: WIKIPEDIA)

A lo largo de la historia ha existido cierto enfrentamiento entre dos escuelas de pensamiento con respecto a la IA: **el enfoque simbólico-lógico (originalmente llamados neats) o top-down y el enfoque basado en datos (originalmente conocidos como scruffies) o bottom-up**. El enfoque simbólico postulaba que para desarrollar máquinas que razonaran era necesario que siguiesen un conjunto de reglas predefinidas y principios de la lógica. Por otra parte, la escuela bottom-up proponía que la inteligencia artificial debería inspirarse en la biología, aprendiendo a partir de la observación y la experiencia, es decir, a partir de datos.

## **En los sistemas de Inteligencia Artificial suele tomarse como referencia la inteligencia humana.**

Del mismo modo que la inteligencia humana es diversa y múltiple, la Inteligencia Artificial es una disciplina con numerosas ramas de conocimiento. La escuela simbólico-lógica, incluye, entre otras, la teoría de juegos, la lógica, la optimización, el razonamiento y la representación del conocimiento, la planificación automática y la teoría del aprendizaje. Por otra parte, en la escuela bottom-up, destacaría: la percepción computacional (una de mis áreas de especialidad, que abarca el procesamiento de imágenes, videos, texto, audio y datos de otro tipo de sensores), el aprendizaje automático estadístico (machine learning, otra de mis áreas de especialidad), el aprendizaje con refuerzo, los métodos de búsqueda (por ejemplo, de texto, imágenes, videos...), los sistemas de agentes, la robótica, el razonamiento con incertidumbre, la colaboración humano-IA, los sistemas de recomendación y personalización y las inteligencias social y emocional computacionales.

Al principio, **el enfoque bottom-up —en el que enmarco mis trabajos—** no tuvo mucho éxito práctico ya que no había disponibilidad de grandes cantidades de datos y capacidades de computación para entrenar los modelos. Por ello, el primer éxito de aplicación práctica de los sistemas de Inteligencia Artificial fue en los años 60 con los sistemas expertos que pertenecen al enfoque simbólico-lógico. De hecho, el periodo entre 1956 y 1974 suele conocerse como la primera etapa dorada de la Inteligencia Artificial. Feigenbaum —uno de los fundadores del departamento de informática de la Universidad de Stanford— lideró el equipo de construyó el primer sistema experto, implementado en LISP, el programa de ordenador desarrollado por McCarthy.

## Expectativas no cumplidas

El periodo entre 1956 y 1974 suele conocerse como la primera etapa dorada de la Inteligencia Artificial, en la que se vivió una época de optimismo con relación a esta disciplina y su impacto. En 1956, después de la famosa convención de Dartmouth, Herbert Simon predijo que “en veinte años, las máquinas serán capaces de hacer el trabajo de una persona”, y en 1970, Marvin Minsky declaró en la revista Life que “dentro de tres a ocho años tendremos una máquina con la inteligencia general de un ser humano”.

Sin embargo, estas expectativas no se cumplieron y entre 1974 y 1980 la Inteligencia Artificial vivió lo que se conoce como su **primer invierno**, que se caracterizó por un declive en el interés hacia la Inteligencia Artificial y su financiación. Además, en 1969 el interés hacia los modelos bottom-up y en particular hacia las redes neuronales decayó significativamente, especialmente tras la publicación del libro Perceptrons por parte de Minsky y Papert, a quien también tuve el honor de conocer durante mis años en MIT. En el libro demostraron que los perceptrones eran muy limitados ya que no podían aprender la función XOR<sup>2</sup> —al no ser linealmente separable— y por tanto no tendrían la capacidad de resolver problemas reales. Los investigadores en Inteligencia Artificial se encontraron con limitaciones y dificultades insalvables en la década de los 70, incluyendo capacidades de computación muy limitadas que hacían imposible el procesamiento de grandes cantidades de datos para entrenar modelos complejos, necesarios para poder abordar problemas reales. No ha sido hasta hace aproximadamente una década cuando hemos podido solventar esta limitación.

A principios de los 80, el interés por la Inteligencia Artificial y su financiación comenzaron a aumentar. De hecho, los primeros sistemas expertos fueron adoptados comercialmente en la década de los 80, de manera que en 1985 el gasto en sistemas de IA en las empresas alcanzó miles de millones de dólares.

2 La función XOR o disyunción exclusiva es un tipo de disyunción lógica que solamente retorna un valor de 1 (verdad) cuando un número impar de entradas es 1 y retorna 0 en el resto de los casos.

**En 1984 nació Cyc3 , el primer esfuerzo científico por implementar en una máquina el razonamiento de sentido común, creando una base de datos masiva que contendría todo el conocimiento sobre el mundo que tiene de media una persona.**

No obstante, fue en ese mismo año, 1984, en el que, durante el congreso de la Asociación Americana de Inteligencia Artificial, Minsky and Schank alertaron de que el entusiasmo y la inversión en la Inteligencia Artificial conducirían a la decepción, de manera que en 1987 comenzó el **segundo invierno** de la Inteligencia Artificial, que alcanzaría su momento más oscuro en 1990.

Sin embargo, la comunidad científica siguió avanzando en las dos escuelas de pensamiento. Uno de los hitos más importantes de la escuela bottom-up y en particular del conexionismo fue la propuesta del uso del **algoritmo de back-propagation** por Rumelhart, Hinton y Williams [11] en 1986 para generar representaciones internas útiles de los datos de entrada en las capas ocultas de las redes neuronales (ver ejemplo en la Figura 5). Aunque Rumelhart et al. no fueron los primeros en publicar un artículo sobre backpropagation, su artículo destaca por la claridad con la que presentan esta idea de manera que consiguieron impactar en la comunidad científica. De hecho, este algoritmo es la base de la gran mayoría de modelos de redes neuronales profundas de hoy en día y permite entrenar redes multi-capa a partir de datos. También cabe destacar el trabajo de Judea Pearl (miembro correspondiente de la Real Academia de Ingeniería) a finales de los 80 en el que incorporó teoría de la probabilidad y teoría de la decisión en la Inteligencia Artificial. Algunos de los métodos nuevos propuestos incluyen las redes Bayesianas, los modelos ocultos de Markov (ambos métodos son clave en mi investigación), la teoría de la información, el modelado estocástico y la optimización. También se desarrollaron los algoritmos evolutivos.

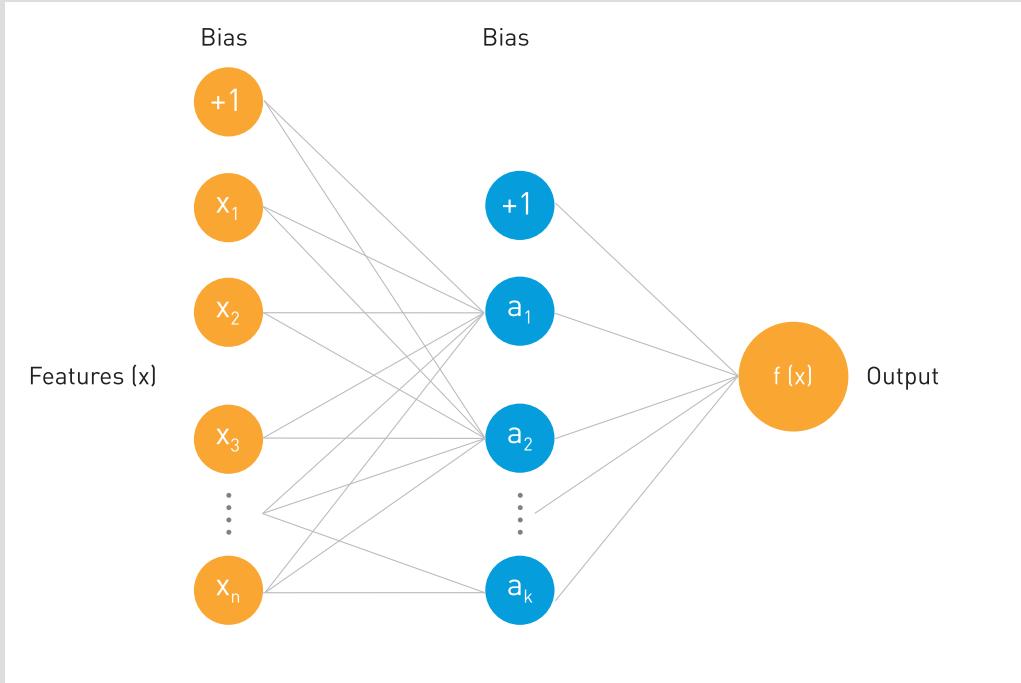


FIG. 5. PERCEPTRÓN MULTI-CAPA CON UNA CAPA OCULTA

Desde mediados de los años 90, precisamente cuando comencé mi doctorado en MIT, hasta hoy en día, y especialmente desde los últimos 10 años, se ha producido un avance muy significativo en las técnicas de aprendizaje estadístico por ordenador basadas en datos (statistical machine learning) que pertenecen al enfoque bottom-up. Mis contribuciones científicas en este campo consisten en el desarrollo y uso de nuevos modelos gráficos dinámicos para modelar distintos comportamientos humanos, como ilustra la Figura 6. Comencé utilizando modelos de Markov ocultos (HMMs) para reconocer en tiempo real expresiones faciales en un sistema llamado LAFTER que recibió el premio a publicación destacada en la revista Pattern Recognition en el año 2000 [12] [1]. Propuse una nueva arquitectura de modelos de Markov llamados modelos de Markov acoplados (coupled HMMs o CHMMs) para reconocer automáticamente interacciones humanas a partir de información captada por videos [3] [13].

Este trabajo ha sido citado extensamente (más de 4.000 citas) e incluido en bibliografías universitarias de visión por ordenador y aprendizaje automático. Otro modelo gráfico novedoso que desarrollé son los modelos ocultos de Markov por capas (layered HMMs), capaces de aprender modelos jerárquicos a partir de datos de diferentes sensores, es decir, en sistemas multi-sensoriales. Ilustré este nuevo método en SEER, un sistema multimodal de reconocimiento en tiempo real de actividades en la oficina, a partir de datos de video, micrófono y teclado y ratón del ordenador. Una de las publicaciones resultado de este proyecto recibió el premio al impacto después de 10 años en el congreso ICMI del ACM [2] y demostramos SEER junto a Bill Gates ante miles de personas en su charla invitada en el congreso IJCAI 2001 en Seattle.

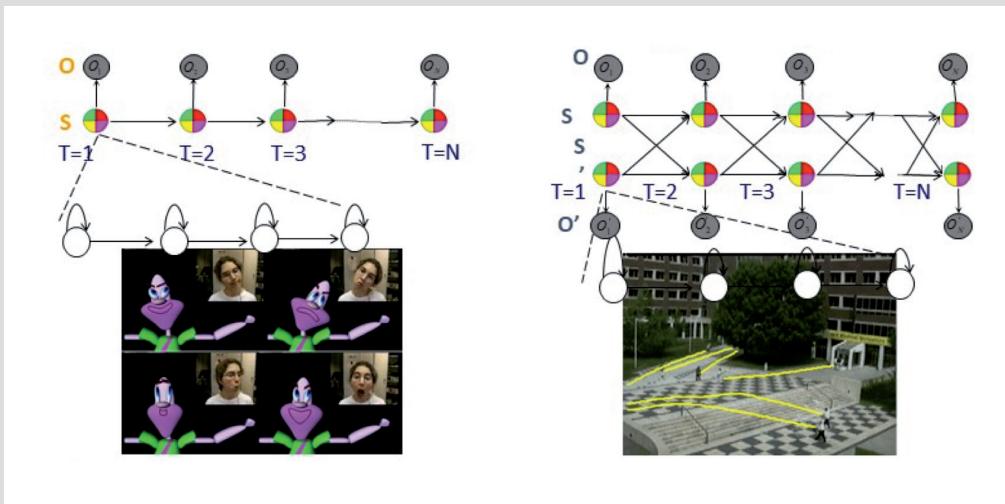
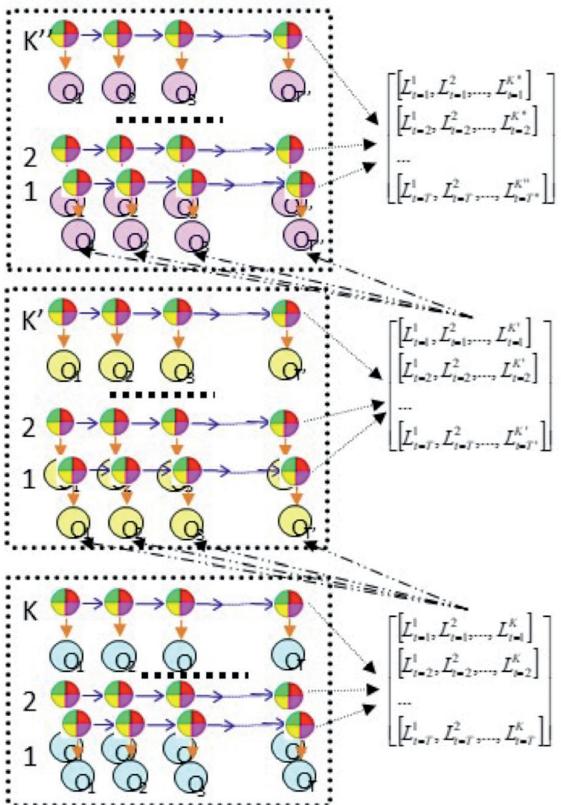


FIG. 6. DIFERENTES MODELOS GRÁFICOS DINÁMICOS PARA RECONOCER EXPRESIONES FACIALES (ARRIBA IZQUIERDA), INTERACCIONES HUMANAS (ARRIBA DERECHA) Y ACTIVIDADES EN LA OFICINA (PÁGINA SIGUIENTE).



Persona  
presente - Otra



Conversación  
cara a cara



Presentación



En el teléfono



Conversación  
distante



Nadie  
alrededor



La disponibilidad de cantidades ingentes de datos —Big Data— como resultado de la actividad humana a través del uso de servicios digitales así como de procesos de digitalización del mundo físico, acompañada por la disponibilidad de procesadores muy potentes a bajo coste y el desarrollo de arquitecturas de redes neuronales profundas y complejas —modelos de deep learning [14] entrenados con dichos datos (ver ejemplos en las Figuras 7 y 8), ha dado lugar a que la Inteligencia Artificial viva hoy en día una “**primavera perpetua**”—en palabras del profesor y amigo de la Universidad de Stanford Andrew Ng, con quien también coincidí en MIT.

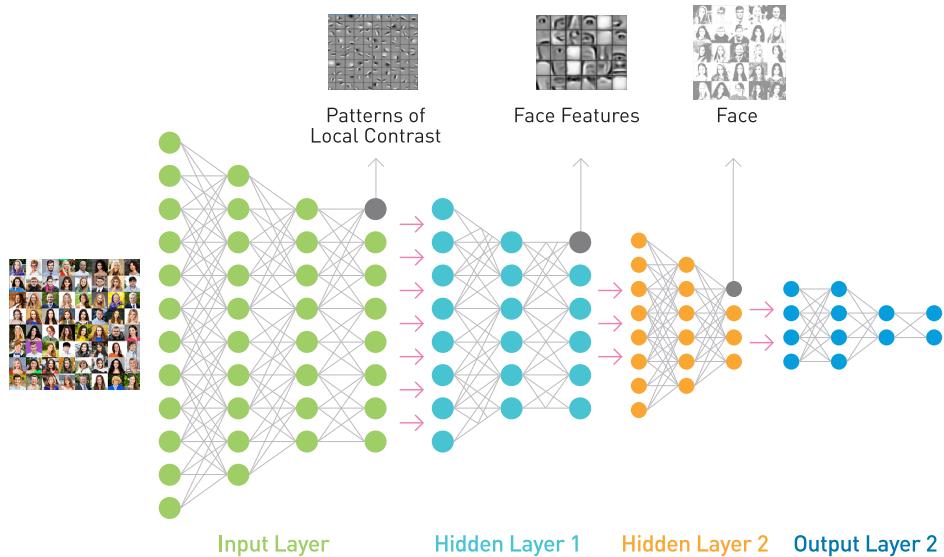
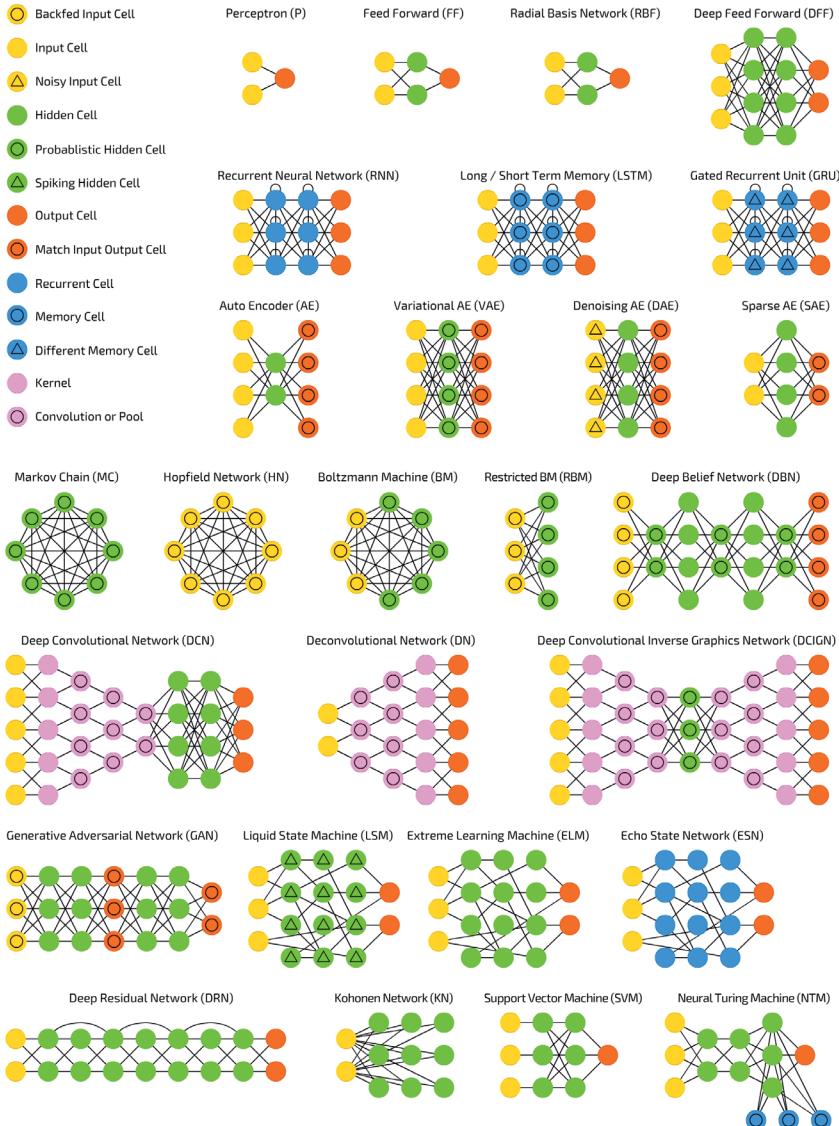


FIG. 7. EJEMPLO DE RED NEURONAL PROFUNDA (DEEP NEURAL NETWORK) PARA EL PROCESAMIENTO Y ANÁLISIS DE CARAS EN IMÁGENES  
(FUENTE: [HTTPS://CDN.EDUREKA.CO/BLOG/WP-CONTENT/UPLOADS/2017/05/DEEP-NEURAL-NETWORK-WHAT-IS-DEEP-LEARNING-EDUREKA.PNG](https://cdn.edureka.co/blog/wp-content/uploads/2017/05/DEEP-NEURAL-NETWORK-WHAT-IS-DEEP-LEARNING-EDUREKA.PNG))

FIG. 8 (PÁGINA SIGUIENTE). EJEMPLOS DE ARQUITECTURAS DE REDES NEURONALES PROFUNDAS (FUENTE: ASIMOVINSTITUTE.ORG FJODOR VAN VEEN)

# A MOSTLY COMPLETE CHART OF NEURAL NETWORKS



En particular, y como puede observarse en la Figura 9, con el éxito de los métodos de Deep learning se ha producido un fuerte resurgir del acercamiento bottom-up y en particular el conexiónismo dentro de la Inteligencia Artificial.

El interés por la Inteligencia Artificial ha experimentado unos niveles exorbitantes, en parte alimentado por la consecución de hitos importantes que hace tan solo tres décadas parecían inalcanzables y que han recibido considerable atención mediática, entre los que destacaría: la victoria de Deep Blue a Gary Kasparov, campeón mundial de ajedrez, en 1997; la conducción automática de 131 millas en un camino en el desierto en 2005 por parte del vehículo autónomo desarrollado por Stanford como parte del DARPA Grand Challenge; la victoria de Watson en el juego de Jeopardy! ante los dos campeones de Jeopardy! del mundo, Brad Rutter y Ken Jennings, en 2011; el desarrollo desde 2011 de asistentes personales en los teléfonos móviles (Siri, Cortana y Google Now) y para el hogar desde el 2015 (Alexa, Google Home) que permiten a sus usuarios utilizar la voz y el lenguaje natural para hacer preguntas y pedir que lleven a cabo acciones de manera automática; la victoria en 2016 de AlphaGo de DeepMind frente a Lee Sedol, uno de los mejores jugadores de Go del mundo; la victoria del programa AI Libratus para jugar al Póker ante sus 4 oponentes humanos —situados entre los mejores jugadores del mundo— con una estadística de victorias impresionante; la capacidad de AlphaZero de no solo vencer al mejor programa del mundo de ajedrez, sino también de aprender a jugar al ajedrez por sí mismo y la superación de los resultados de los humanos en la prueba de lectura y comprensión lectora de Stanford en un conjunto de cien mil preguntas por el sistema de procesamiento de lenguaje desarrollado por Alibaba, ambos en 2018.

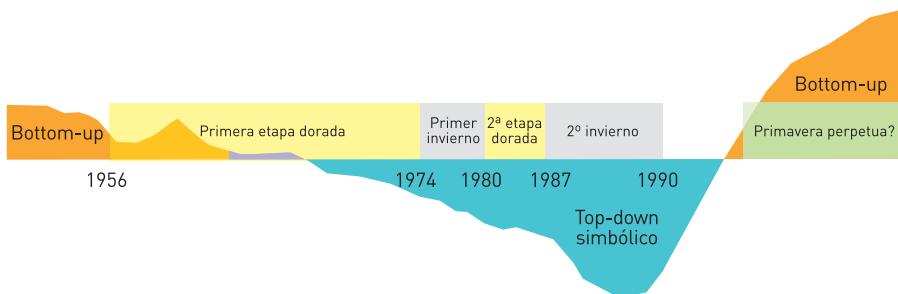
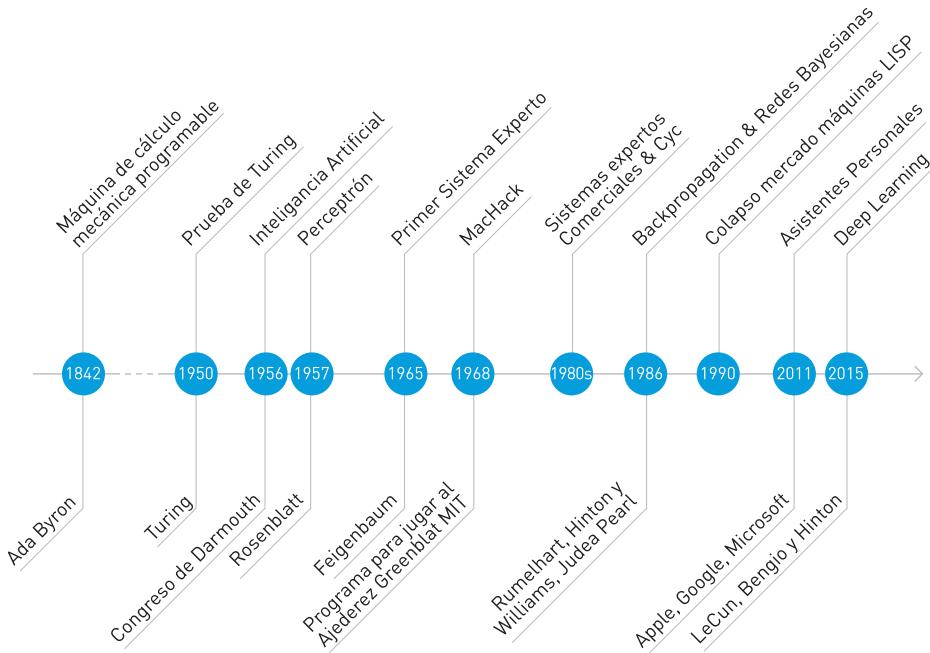


FIG. 9. CRONOLOGÍA RESUMIDA DE LA EVOLUCIÓN DE LA INTELIGENCIA ARTIFICIAL. ALTERNANCIA ENTRE LA PREPONDERANCIA DE MODELOS BOTTOM-UP (NARANJA) VS MODELOS TOP-DOWN (AZUL) A LO LARGO DE LA HISTORIA, CON CLARO AUGE DE LOS MODELOS BOTTOM-UP Y ESPECIALMENTE EL CONEXIONISMO EN LOS ÚLTIMOS AÑOS

Mas allá de estos hitos, la Inteligencia Artificial ya ocupa un lugar importante en multitud de esferas de nuestra vida. Está presente en los sistemas de búsqueda y recomendación de información/contenido/productos/amigos que utilizamos en nuestro día a día (por ejemplo, Netflix, Spotify, Facebook, servicios de noticias), en las aplicaciones de cámara del móvil que detectan automáticamente las caras en las fotos, en los asistentes personales de nuestros móviles usando la voz o el texto (como Siri, Cortana, Google Now, Alexa o Google Home, chatbots conversacionales) o en las ciudades inteligentes para, por ejemplo, predecir el tráfico. También forma parte de los sistemas para la toma de decisiones en los mercados financieros y en las empresas (por ejemplo, compraventa de acciones, adjudicación de créditos, seguros, fijación de tarifas), en soluciones médicas (por ejemplo, los sistemas de elaboración automática de diagnósticos clínicos y de análisis automático de imágenes médicas, historiales clínicos y ADN), en los sistemas de toma de decisión de las Administraciones Públicas (por ejemplo, sistemas de vigilancia, de soporte a decisiones judiciales o de clasificación y jerarquización del alumnado público), en los procesos productivos (por ejemplo, robots industriales, sistemas de planificación y predicción de la demanda/producción...), en los vehículos autónomos, en la predicción meteorológica y los modelos físicos, y en sistemas de seguridad y defensa (por ejemplo, armas autónomas, control fronterizo de personas y adjudicación de visados).

A lo largo de mi trayectoria, he contribuido a la personalización y sistemas de recomendación con modelos que incorporan el contexto [15], que optimizan la selección de resultados en forma de una lista [16] o que incluyen las opiniones de expertos [17]. Entre otros, también he desarrollado MobiScore, un sistema para la inferencia de riesgo crediticio en economías en desarrollo a partir de patrones de uso del móvil [6], un proyecto para detectar automáticamente puntos calientes (hotspots) de personas [18] o de crímenes [7] en las ciudades mediante el análisis agregado de datos de la red de telefonía móvil y un método pionero para segmentar automáticamente el uso de una ciudad a través del análisis de los patrones de uso de los sistemas de bicicletas compartidas [19].

No podemos obviar la contribución de los avances en el hardware —y en particular sistemas de procesamiento y almacenamiento a gran escala, distribuidos y en paralelo— para el desarrollo y aplicación práctica de la IA, así como la consecución de los hitos anteriormente descritos. El entrenamiento a partir de cantidades ingentes de datos de modelos complejos de deep learning con un consumo energético y en un tiempo razonables no sería posible sin la existencia de nuevos procesadores, en muchos casos optimizados para el entrenamiento de algoritmos de aprendizaje profundo. En los últimos años hemos pasado de utilizar procesadores de propósito general (CPUs y GPUs o graphics processing units) al uso de procesadores especializados y optimizados para este tipo de modelos (FPGAs o field-programmable gate arrays y ASICs o application-specific integrated circuits como la TPU desarrollada por Google).

Las Figuras 10 y 11 ilustran la evolución en la capacidad de computación. La Figura 10 muestra la famosa Ley de Moore desde 1971, según la cual el número de transistores que podemos integrar en un circuito se duplica cada 1-1,5 años por el mismo precio. En la parte de la izquierda se muestra la Ley de Moore en CPUs y en la derecha en GPUs. La Figura 11 ilustra donde se sitúan los distintos tipos de procesadores en función de su flexibilidad y facilidad de uso versus su eficiencia energética y rendimiento. El lector interesado puede encontrar un resumen de los principales procesadores utilizados para el aprendizaje automático en [20] utilizados para aprendizaje automático en [20].

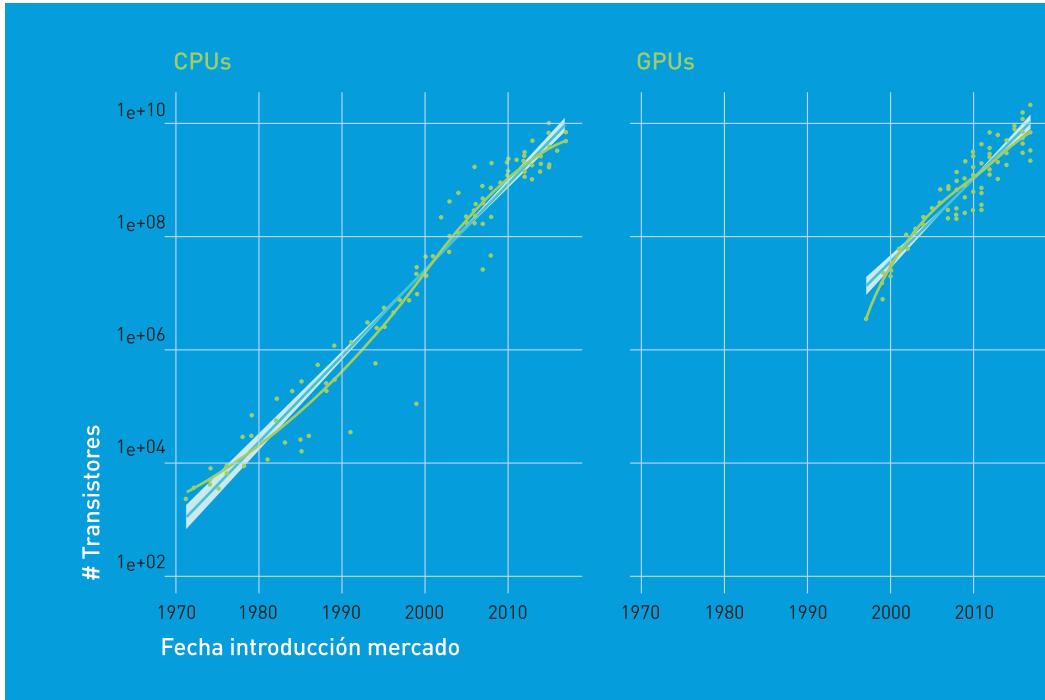
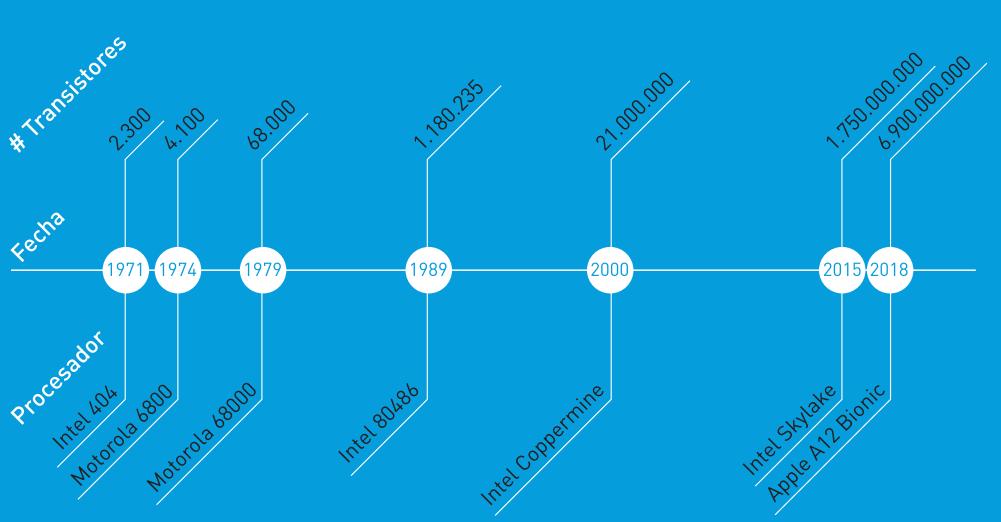


FIG 10. LEY DE MOORE DESDE 1971 PARA LAS CPUS (ARRIBA) Y GPUS (PÁGINA SIGUIENTE).  
(FUENTE: FIGURAS PROPIAS A PARTIR DE [HTTP://COULMONT.COM](http://COULMONT.COM) A PARTIR DE DATOS DE WIKI-PEDIA)



Finalmente, no podemos obviar el consumo energético necesario para poder transmitir y procesar las cantidades ingentes de datos existentes hoy en día. En un estudio reciente [21], Andrae estima que, si no conseguimos una mejora dramática en la eficiencia energética de los procesadores utilizados, la industria TIC podría consumir el 20% de toda la electricidad del planeta y emitir hasta un 5,5% de las emisiones de CO<sub>2</sub> del planeta en el año 2025. La predicción de demanda de electricidad iría de 200-300 TWh de electricidad al año a 1.200-3.000 TWh en 2025. Los centros de datos podrían producir 1,9Gt (o un 3,2% del total) de las emisiones de carbón del planeta.

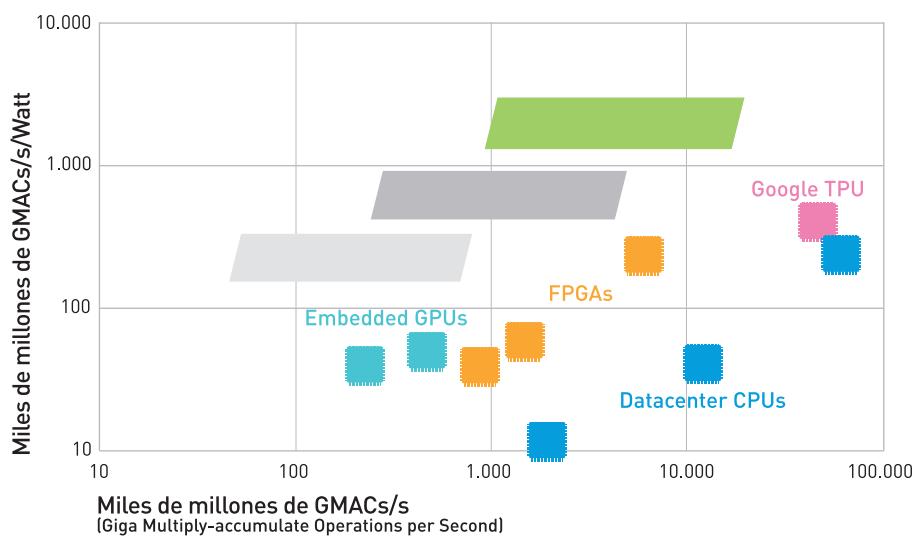
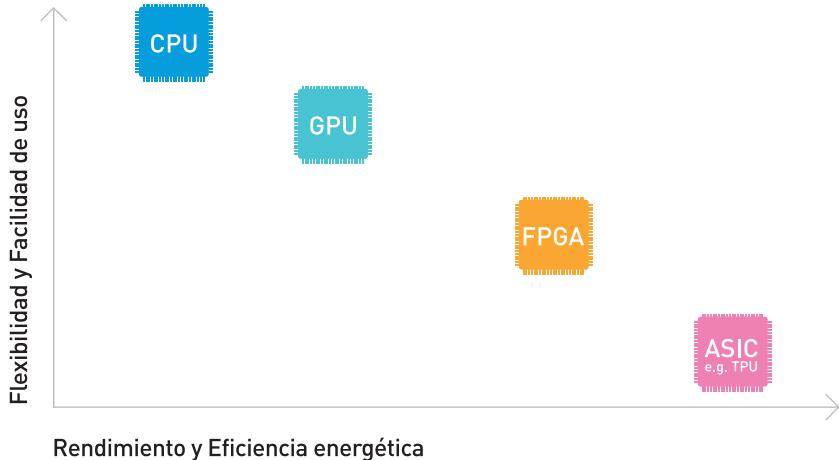


FIG 11. ARRIBA. ILUSTRACIÓN DE LOS DISTINTOS TIPOS DE PROCESADORES UTILIZADOS EN MODELOS DE APRENDIZAJE PROFUNDO;  
SOBRE ESTAS LÍNEAS. RENDIMIENTO VS EFICIENCIA DE DISTINTOS TIPOS DE PROCESADORES  
(FUENTE: FIGURAS PROPIAS A PARTIR DE DATOS DE CHRIS ROWEN, COGNITE VENTURES)



Respecto a su nivel de competencia, los sistemas de Inteligencia Artificial suelen dividirse en tres tipos:

### **1. Sistemas con IA específica,**

que es el tipo de Inteligencia Artificial que tenemos hoy en día. Son sistemas capaces de realizar una tarea concreta (por ejemplo, reconocer el habla, reconocer imágenes, procesar texto...) incluso mejor que un humano, pero solamente esa tarea;

### **2. Sistemas con IA general,**

que sería la aspiración de la Inteligencia Artificial ya que son sistemas que exhiben una inteligencia similar a la humana: múltiple, adaptable, flexible, eficiente, incremental... Estamos muy lejos de llegar a este tipo de sistemas;

### **3. Sistemas con súper-inteligencia,**

término un tanto controvertido que hace referencia al desarrollo de sistemas que tendrían una inteligencia superior a la humana, tal y como propone el filósofo británico Bostrom [22].

Hoy en día disponemos de sistemas de Inteligencia Artificial específica, es decir, capaces de realizar automáticamente y autónomamente una tarea concreta, pero sólo esa tarea. Por ejemplo, aunque un algoritmo diseñado para jugar al ajedrez juegue mejor que el mejor de los humanos, es incapaz de hacer cualquier otra tarea. De hecho, tampoco "sabe" qué es el ajedrez y tendría dificultades para poder jugar si hiciésemos una variación a las reglas del juego. Por ello, los sistemas actuales manifiestan un tipo limitado de inteligencia ya que son incapaces, entre otras cosas, de generalizar y extender sus niveles de competencia en una cierta tarea a otros ámbitos de manera automática, como haría un humano.



El éxito reciente de los sistemas de Inteligencia Artificial —como he ilustrado anteriormente— quizás este desviando la atención respecto a problemas fundamentales de la Inteligencia Artificial actual que todavía están por resolver. Entre otros: (a) la necesidad de incorporar semántica y razonamiento en los sistemas de procesamiento del lenguaje natural; (b) la necesidad de desarrollar representaciones de la incertidumbre que sean tratables computacionalmente; (c) la importancia de desarrollar sistemas que puedan formular, adaptarse y perseguir objetivos a largo plazo; (d) la importancia de representar e inferir la causalidad; (e) la necesidad de incorporar la información de contexto en los modelos; (f) la capacidad para aprender constante, incremental y asociativamente; y (g) la importancia de que los modelos estén dotados de robustez, de manera que no fallen estrepitosamente cuando se cambian ciertas características en los datos de entrada, como sucede hoy en día.

Algunos de estos problemas derivan de tres limitaciones básicas en los mecanismos de aprendizaje automático disponibles actualmente que los diferencian de los sistemas inteligentes biológicos, como explica Jeff Hawkins en un artículo para IEEE Spectrum [23] y que resumo en los siguientes párrafos.

**En primer lugar, los sistemas de aprendizaje biológicos (cerebros) son capaces de aprender rápidamente;** un par de observaciones o de experiencias táctiles suelen ser suficientes para aprender algo nuevo, a diferencia de los millones de ejemplos que necesitan los sistemas de Inteligencia Artificial actuales. Asimismo, los sistemas biológicos aprenden de manera incremental, es decir, podemos agregar conocimiento nuevo sin tener que volver a aprender todo desde cero o perder conocimiento anterior, y de manera continua, es decir, nunca dejamos de aprender conforme interaccionamos con el mundo físico en el que vivimos. El aprendizaje rápido, incremental y constante es un elemento esencial que permite a los sistemas biológicos inteligentes adaptarse a un entorno cambiante y sobrevivir. **La neurona** es un elemento clave en el aprendizaje biológico y la complejidad de las neuronas biológicas y sus conexiones es lo que dotan al cerebro de la capacidad para aprender. Hoy sabemos que el cerebro tiene plasticidad y que constantemente se están creando nuevas neuronas (neurogénesis) y sinapsis (sinaptogénesis). De hecho, hasta un 40% de las sinapsis en una neurona podrían ser sustituidas cada día con nuevas sinapsis, que dan lugar a nuevas conexiones entre neuronas. Los sistemas

de aprendizaje artificiales no tendrían que reproducir exactamente cómo funcionan las neuronas biológicas, pero esta capacidad para disponer de un aprendizaje rápido, incremental y constante, caracterizado por la destrucción y creación de sinapsis es esencial.

**En segundo lugar, el cerebro utiliza lo que se conoce como representaciones distribuidas no densas o dispersas (sparse distributed representations o SDRs)** [24]. Se llaman dispersas porque sólo un conjunto reducido de neuronas está activo en cada momento del tiempo. Qué neuronas están activas cambia de un instante a otro, en función de lo que haga el ser vivo, pero el conjunto de neuronas que está activo es pequeño. Este tipo de representación es robusta a errores y al ruido. Además, gozan de dos propiedades interesantes: la propiedad del solape, que permite detectar rápidamente si dos percepciones son idénticas o diferentes; y la propiedad de la unión, que permite al cerebro mantener varias representaciones en paralelo. Por ejemplo, si sentimos que hay un animal que se mueve entre unos arbustos, pero no hemos podido verlo con claridad, podría ser un conejo, una ardilla o una rata. Dado que las representaciones en el cerebro son poco densas, nuestro cerebro puede activar tres SDRs al mismo tiempo (la del conejo, la ardilla y la rata) sin que haya interferencia entre ellas. Esta propiedad de poder unir SDRs es lo que permite al cerebro gestionar, operar y tomar decisiones con incertidumbre.

**En tercer lugar, el aprendizaje tiene “cuerpo” (embodiment), de manera que nuestro cerebro recibe la información de los distintos sentidos, información que cambia según nos movemos y actuamos en nuestro entorno.** Un sistema de Inteligencia Artificial no tiene por qué tener un cuerpo físico, pero sí la capacidad de actuar sobre su entorno (físico y virtual) y recibir feedback en función de sus acciones. Los sistemas de aprendizaje con refuerzo (reinforcement learning) hacen algo similar y son instrumentales en la consecución de algunos de los hitos previamente descritos, como es el caso del sistema AlphaZero. Además, el cerebro es capaz de integrar la información captada por los distintos sentidos y por el sistema motor, para poder no sólo procesar, reconocer y decidir en base a lo percibido, sino también actuar. Esta integración sensorial-motora es básica en el funcionamiento del cerebro y probablemente debería también serlo de los sistemas de Inteligencia Artificial.

**Finalmente, en el cerebro, la información sensorial es procesada por un sistema jerárquico de manera que conforme la información va pasando de un nivel a otro, se van calculando características cada vez más complejas sobre lo que se está percibiendo.** Los modelos de aprendizaje profundo (**deep-learning**) también utilizan jerarquías, pero mucho más profundas, con decenas o una centena de niveles y centenas o miles de millones de parámetros, mientras que el cerebro necesita pocos niveles. Asimismo, como ya he comentado, las redes neuronales profundas necesitan millones de observaciones para poder aprender un patrón, mientras que el cerebro necesita pocos ejemplos. **El cerebro tiene una capacidad para aprender que es mucho más eficiente que la de los modelos computacionales de hoy en día.**

Sin embargo, a pesar de estas limitaciones, la Inteligencia Artificial ya está teniendo un gran impacto en la sociedad —como he descrito anteriormente— y forma parte integral de la Cuarta Revolución Industrial, en la que estamos inmersos, como explico a continuación.



En los últimos tres siglos, hemos vivido cuatro revoluciones industriales, ilustradas en la Figura 12. La **Primera Revolución Industrial** tuvo lugar entre los siglos XVIII y XIX en Europa y América y se corresponde con el momento histórico en que sociedades que eran mayoritariamente agrarias y rurales se convirtieron en industriales y urbanas. El principal impulsor de esta revolución fue la invención de la máquina de vapor, junto con el desarrollo de las industrias textil y metalúrgica.

La **Segunda Revolución Industrial** —conocida como la Revolución Tecnológica— ocurrió justo antes de la Primera Guerra Mundial, entre 1870 y 1914, y se corresponde con un crecimiento de las industrias anteriores y el desarrollo de nuevas industrias como la industria del acero, el petróleo y la electricidad. Los avances tecnológicos más importantes de esta revolución incluyen el teléfono, la bombilla, el fonógrafo y el motor de combustión interna.

La **Tercera Revolución Industrial** se conoce como la Revolución Digital y hace referencia a la transición de dispositivos mecánicos y analógicos al uso de tecnologías digitales. Comenzó en los años 80 y continua hoy en día. Los avances tecnológicos clave en esta Tercera Revolución Industrial incluyen los ordenadores personales, internet y el desarrollo de otras tecnologías de la información y las comunicaciones (TICs).

Finalmente, la **Cuarta Revolución Industrial** se apoya en avances de la Revolución Digital, pero incorpora la ubiquidad de la tecnología digital tanto en nuestra sociedad como en nuestro cuerpo y la unión creciente entre el mundo físico y el mundo digital. Los avances tecnológicos que hacen que esta nueva revolución sea posible incluyen a la robótica, la Inteligencia Artificial, el Big Data, la nanotecnología, la biotecnología, el internet de las cosas, los vehículos autónomos, las impresoras en tres dimensiones y la computación cuántica. El término fue presentado y reconocido globalmente durante el Foro Económico Mundial en 2016 por su fundador, Klaus Schwab<sup>4</sup>.

4 <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>



FIG 12. DIAGRAMA DE LAS CUATRO REVOLUCIONES DESDE EL S. XVIII

La Inteligencia Artificial tiene un conjunto de características que contribuyen a que sea uno de los elementos clave en esta Cuarta Revolución Industrial, incluyendo:

## 1. Transversalidad e invisibilidad

Las técnicas de Inteligencia Artificial pueden utilizarse en un sinfín de aplicaciones en áreas como la biología, la física, la medicina, la química, la energía, el transporte, la educación, los sistemas de producción, la logística y el transporte, los servicios digitales y la prestación de servicios públicos y privados. Además, la gran mayoría de sistemas de Inteligencia Artificial que se utilizan hoy en día son invisibles, es decir, consisten en software en el corazón de los sistemas y servicios inteligentes cotidianos. Estas dos propiedades —transversalidad e invisibilidad— posicionan a la Inteligencia Artificial en el corazón de la Cuarta Revolución Industrial, con un papel similar al que jugó la electricidad en la Segunda Revolución Industrial<sup>5</sup>.

5 <https://medium.com/syncedreview/artificial-intelligence-is-the-new-electricity-andrew-ng-cc132ea6264>

## 2. Complejidad, escalabilidad y actualización constante

Los sistemas actuales de Inteligencia Artificial basados en modelos de aprendizaje profundo son complejos, con cientos de capas de neuronas y millones de parámetros. Esta complejidad dificulta la capacidad para interpretar los modelos que en ciertos casos de uso —por ejemplo, en la medicina o en la educación— es condición necesaria para poder aplicar un sistema de IA. Al mismo tiempo, esta complejidad permite que los sistemas de Inteligencia Artificial puedan procesar cantidades ingentes de datos —que de otra manera sería inviable— y realizar tareas con niveles de competencia superiores a los de los humanos. Es decir, dotan a los sistemas de IA de gran escalabilidad. En muchos de los casos de uso del Big Data, sólo podemos extraer conocimiento y valor de tales cantidades de datos a través del uso de sistemas de Inteligencia Artificial, ya que los métodos tradicionales no escalan a volúmenes de datos tan grandes que, además, son variados, no estructurados y generados a gran velocidad. Asimismo, los sistemas de Inteligencia Artificial son altamente escalables al consistir fundamentalmente en software, que, además, puede estar conectado con miles o millones de otros sistemas de IA, dando lugar a una red colectiva de inteligencia artificial. Esta escalabilidad, combinada con la capacidad de actualizar el software de manera masiva, permitiría en un periodo corto de tiempo alcanzar y tener impacto en cientos o incluso miles de millones de personas. Podemos imaginar escenarios sin precedentes en la historia de la humanidad. Escenarios donde, por ejemplo, una red de sistemas de IA podría muy rápidamente incorporar los últimos métodos computacionales en el diagnóstico de una enfermedad y desplegarlos a toda la población del planeta. El equivalente analógico, que consistiría en una incorporación casi instantánea del conocimiento a todos los médicos del planeta, es inviable.

## 3. Habilidad para predecir

Los sistemas de Inteligencia Artificial pueden utilizarse para la toma de decisiones automáticas y para predecir situaciones futuras. De hecho, la aspiración es que las decisiones algorítmicas basadas en IA entrenada con datos carezcan de las limitaciones de las decisiones humanas (por ejemplo, conflictos de interés, sesgos, intereses propios, corrupción...) y por tanto sean más justas y objetivas. Sin embargo, esto no es necesariamente así si no

tenemos en cuenta las limitaciones de las decisiones algorítmicas basadas en datos, como explicaré posteriormente.

Los sistemas actuales de Inteligencia Artificial necesitan acceder a grandes cantidades de datos y de computación, así como experto/as en la materia para diseñar y entrenar dichos sistemas. Existe, desgraciadamente, una situación de asimetría de manera que sólo una minoría tiene acceso a dichos datos, capacidad de computación, conocimiento y experiencia, pudiendo beneficiarse plenamente de la Inteligencia Artificial; mientras que una mayoría carece de dichas capacidades y por tanto es una mera usuaria de dicha tecnología. Esta característica plantea un reto importante, ya que deberíamos minimizar la asimetría para garantizar que el impacto de la Inteligencia Artificial sea positivo en toda la sociedad y no solo en un subconjunto de esta.

Además, las técnicas del estado del arte de la Inteligencia Artificial son capaces de generar contenido sintético (fotos, texto, audio y videos) indistinguible del contenido real. Esta capacidad para inventar contenido está transformando la comunicación, la difusión de la información y la formación de la opinión pública. Otorga poder sin precedentes a quienes tengan la capacidad para generar y difundir dicho contenido artificial [25]. De nuevo, esta característica plantea un reto fundamental que deberíamos abordar.

A pesar de estas limitaciones, la presencia de la IA en nuestras vidas y su capacidad para tener impacto positivo en la sociedad son innegables. **Por ello, las grandes potencias mundiales —tanto empresas como gobiernos— han comprendido que el liderazgo en la Inteligencia Artificial conllevará un liderazgo no solo a nivel económico sino también político y social, dado su papel central en la Cuarta Revolución Industrial.** De hecho, en los últimos 2 años, los gobiernos de una veintena de países<sup>6</sup> —incluyendo EEUU, China, Canadá, Francia, Taiwán, Singapur, México, Suecia, India, Australia, Finlandia— han elaborado estrategias nacionales sobre la Inteligencia Artificial, tal y como refleja la Figura 13. A nivel europeo, la Comisión Europea publicó en abril del 2018 una comunicación sobre la Inteligencia Artificial<sup>7</sup> y se espera que antes de que finalice el 2018 publiquen la

6 <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>

7 <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>

estrategia europea sobre Inteligencia Artificial. A nivel nacional, soy miembro del comité de 9 expertos nombrados por el Secretario de Estado para el Avance Digital<sup>8</sup> cuyo cometido es la elaboración de un libro blanco sobre el Big Data y la IA con recomendaciones estratégicas para el Gobierno de España. El libro debería ser publicado antes del verano de 2019.

Como consecuencia de algunas de las características anteriores, el impacto de la IA no estará necesariamente distribuido de manera homogénea o justa en la sociedad. En consecuencia, deberíamos abordar con urgencia las limitaciones de los sistemas existentes de Inteligencia Artificial para asegurarnos de que las oportunidades efectivamente tienen impacto social positivo. En las siguientes secciones, describiré brevemente cuatro dimensiones fundamentales en el contexto del desarrollo de la IA: las dimensiones laboral y económica, social y ética.

## 2017



## 2018



FIG. 13. LÍNEA TEMPORAL DE LA PUBLICACIÓN DE ESTRATEGIAS NACIONALES SOBRE INTELIGENCIA ARTIFICIAL. (FUENTE: FIGURA PROPIA A PARTIR DE DATOS DE TIM DUTTON)

8 Anteriormente llamada Secretaría de Estado de Agenda Digital

## 5. El impacto laboral y económico de la Inteligencia Artificial

El progreso tecnológico asociado a la Cuarta Revolución Industrial está polarizando el mercado laboral. Por una parte, aparecen nuevos trabajos altamente remunerados y que requieren especialización en áreas tecnológicas (por ejemplo, científico/a de datos, ...). Por otra parte, desaparecen otro tipo de trabajos y profesiones al verse automatizadas parcial o totalmente (por ejemplo, taxistas/transportistas, cajeros/as, agentes de viaje...). En particular, **todo lo que pueda automatizarse o sustituirse con el uso de tecnología, será automatizado total o parcialmente**. En consecuencia, la demanda laboral está experimentando un sesgo en beneficio de habilidades profesionales especializadas, y en perjuicio de habilidades u ocupaciones rutinarias y mecánicas. Esta tendencia pronostica un cambio completo en la estructura ocupacional que probablemente conlleve riesgos para la sociedad si no somos capaces de adaptarnos a dicho cambio.

En España, la Organización para la Cooperación y el Desarrollo Económico (OCDE) cifra en el 12% los empleos que pueden estar en riesgo por la automatización [26].

Al mismo tiempo, el desarrollo de tecnologías disruptivas con capacidad para transformar la sociedad ha llevado históricamente una generación de empleo. Según un estudio de McKinsey<sup>9</sup>, un tercio de los nuevos puestos de trabajo creados en EE.UU. en los últimos 25 años pertenecen a disciplinas que no existían anteriormente, en áreas como las tecnologías de la información, la fabricación de hardware, la creación de aplicaciones móviles o la gestión de sistemas tecnológicos. En la próxima década, el empleo se concentrará en funciones cualificadas y con mayor aportación de valor, llegando a tasas de paro inferiores al 3,5% para estos perfiles, frente a una tasa de paro de un 20% para profesiones que requieren una baja cualificación. A nivel mundial, un estudio reciente del Foro Económico Mundial<sup>10</sup> estima que habrá un crecimiento neto de 58 millones de puestos de trabajo en 2022 como consecuencia de la IA.

9 [https://www.mckinsey.com/featured-insights/employment-and-growth/technology-jobs-and-the-future-of-work#section\\_1](https://www.mckinsey.com/featured-insights/employment-and-growth/technology-jobs-and-the-future-of-work#section_1)

10 [http://reports.weforum.org/future-of-jobs-2018/?doing\\_wp\\_cron=1537554858.4296801090240478515625](http://reports.weforum.org/future-of-jobs-2018/?doing_wp_cron=1537554858.4296801090240478515625)

## **La Comisión Europea<sup>11</sup> anticipa una necesidad de más de 900.000 nuevos puestos de trabajo tecnológicos en el corto plazo que seremos incapaces de cubrir si no transformamos nuestros programas educativos.**

En el contexto nacional, el informe EPYCE 2017<sup>12</sup> recoge que el perfil más demandado en España actualmente es el de ingeniero/a informático, seguido por perfiles relacionados con el Big Data. Además, el informe predice que un 49% de las profesiones más demandadas en el futuro en España serán en los ámbitos de la ingeniería y la tecnología. Sin embargo, ¿estamos preparados como sociedad para suplir tal demanda laboral de nuevos puestos de trabajo derivados de la IA? Considero que no.

Por ello, es de vital importancia que invirtamos en la formación de profesionales cuyo trabajo vaya a verse afectado por el desarrollo de la Inteligencia Artificial de manera que puedan seguir contribuyendo a la sociedad en nuevos puestos de trabajo. Estamos progresando hacia un modelo de aprendizaje continuo a lo largo de la vida (**lifelong learning**) y a la desaparición de una única carrera profesional durante la vida laboral de una persona. Gracias a la tecnología disponemos de sistemas de formación continua online (por ejemplo, los cursos masivos abiertos en línea o MOOCs) que ofrecen, de manera escalable y económica, oportunidades de aprendizaje a cualquier persona desde cualquier lugar. Deberíamos asegurarnos de que los profesionales tengan la oportunidad de aprender tecnologías emergentes en sus ámbitos de actividad y así continuar siendo relevantes en sus áreas de competencia, incluso —y especialmente— si estas áreas se ven afectadas por la automatización.

Esta necesidad de aprendizaje constante, consecuencia del cambio también constante causado por el progreso tecnológico, puede ser difícil de gestionar desde un punto de vista emocional. Como bien sabemos, los seres humanos solemos ser resistentes al cambio, especialmente conforme envejecemos.

11 [https://ec.europa.eu/commission/commissioners/2014-2019/ansip/blog/digital-skills-jobs-and-need-get-more-europeans-online\\_en](https://ec.europa.eu/commission/commissioners/2014-2019/ansip/blog/digital-skills-jobs-and-need-get-more-europeans-online_en)

12 [http://marketing.eae.es/prensa/\\_EPyCE2017.pdf](http://marketing.eae.es/prensa/_EPyCE2017.pdf)

Por ello, deberíamos contemplar la posibilidad de la existencia de un colectivo social incapaz de adaptarse a esta necesidad constante de aprendizaje y por tanto carente de las herramientas necesarias para contribuir a la sociedad del mañana. En palabras del historiador Yuval Noah Harari, una “**clase inútil**” [27].

**Algunas de las potenciales soluciones que se han propuesto son la creación de un salario básico universal o la provisión universal por parte de los gobiernos de las necesidades básicas gratuitamente.**

Desde un punto de vista económico, en 2018 el mercado de productos, hardware y software relacionados con la Inteligencia Artificial se espera que supere los 4 mil millones de dólares, según un estudio de Statista<sup>13</sup> (ver Figura 14). Con un crecimiento sostenido en el tiempo, un estudio macroeconómico de PwC<sup>14</sup> estima que la Inteligencia Artificial generará en el año 2030 más de 15 billones (millones de millones) de dólares a nivel mundial, con impacto en todos los ámbitos de actividad, tanto en el sector público como privado. Este impacto económico es debido tanto al aumento de la productividad —fruto de los procesos de automatización y asistencia a la fuerza laboral con sistemas de Inteligencia Artificial— como a una demanda creciente por parte de los consumidores de productos y servicios enriquecidos con IA. Geográficamente, Norte América y China experimentarán los mayores beneficios dada la concentración de la investigación, la innovación y el desarrollo de la Inteligencia Artificial en estas dos regiones.

13 <https://www.statista.com/statistics/607716/worldwide-artificial-intelligence-market-revenues/>

14 <https://www.pwc.co.uk/economic-services/assets/macroeconomic-impact-of-ai-technological-report-feb-18.pdf>

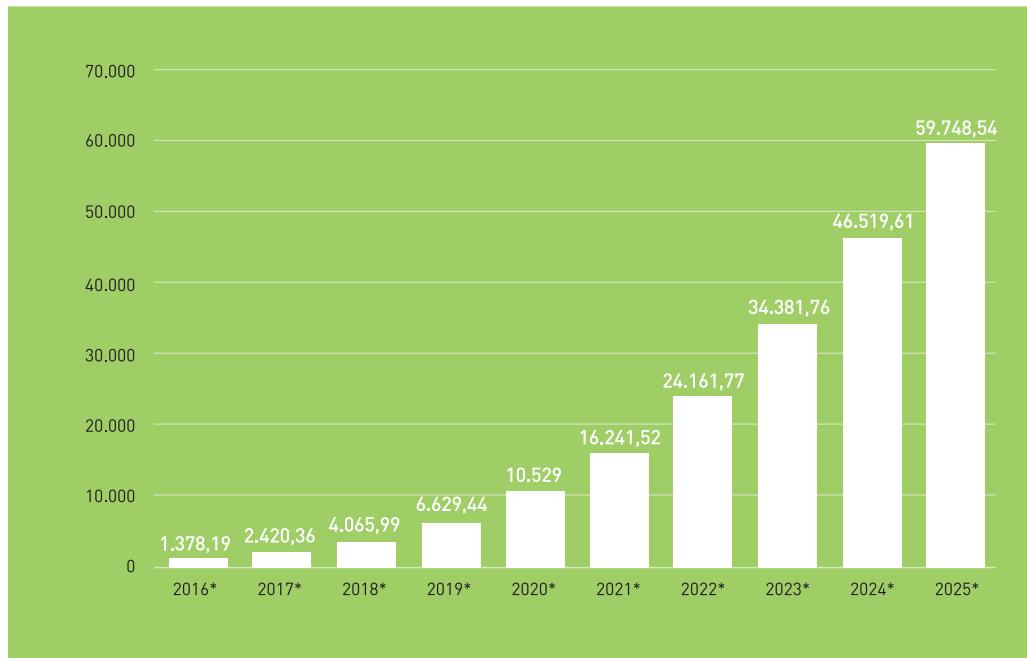


FIG 14. ESTIMACIÓN DEL TAMAÑO DEL MERCADO DE LA INTELIGENCIA ARTIFICIAL EN MILLONES DE DÓLARES AMERICANOS. (FUENTE: STATISTA)

Con respecto a la innovación en IA, un estudio reciente de Asgard y Roland Berger<sup>15</sup> analiza la distribución de startups de IA a nivel mundial. EE.UU. es el líder mundial con un 40% de todas las startups de IA analizadas por este informe (~1400 startups de IA, de las cuales ~600 están en San Francisco), seguidos de China e Israel. A nivel europeo, Londres es la segunda ciudad del mundo en número de startups de IA, seguida de París en décima posición. Finalmente, un informe de Accenture y Frontier Economics<sup>16</sup> estima el crecimiento económico que diferentes países (incluyendo España) podrían tener si invirtiesen en IA y consiguiesen incorporar sus beneficios en la economía, comparado con el crecimiento económico de base, sin el uso de IA. En el caso

15 “The global AI landscape” <https://asgard.vc/global-ai/>

16 <https://www.accenture.com/us-en/insight-ai-industry-growth>

de España, la predicción es de un crecimiento adicional del PIB en un 0.8% en el año 2035 gracias a la IA.

En el contexto español, destacaría las oportunidades que existen en la salud, el transporte, la energía, la agricultura, el turismo, el comercio electrónico, la banca y la administración pública.

## **6. El potencial de la Inteligencia Artificial para mejorar la sociedad**

La sociedad debería verse enriquecida por la Inteligencia Artificial. Es indudable que la Inteligencia Artificial está contribuyendo y contribuirá al crecimiento económico, como he explicado anteriormente. Permitirá que tengamos una medicina de precisión (personalizada, preventiva y predictiva), una educación personalizada y permanente, ciudades inteligentes, una gestión más eficiente de los recursos y una toma de decisiones más justas, transparentes y basadas en evidencia. Asimismo, este impacto no vendrá exento de cambios sociales profundos, incluyendo una transformación del mercado laboral, descrita en la sección anterior.

El desarrollo e implantación de una Inteligencia Artificial centrada en las personas debería resultar en un empoderamiento de la sociedad. Una condición necesaria para este empoderamiento es el conocimiento.

Para ello hemos de invertir tanto en educación formal como informal. Si no lo hacemos, será muy difícil, si no imposible que seamos capaces como sociedad de tomar decisiones sobre tecnologías que no entendemos y que en consecuencia frecuentemente tememos. Coincido plenamente con **las palabras de Marie Curie, “nada en la vida debería temerse, sino entenderse. Ahora es momento de entender más para así temer menos”**.

Desde un punto de vista formal, en las etapas de educación obligatoria (primaria y secundaria) de muchos países del mundo —entre los que desgraciadamente no se encuentra España— ya se ha incorporado una asignatura troncal de Pensamiento Computacional<sup>17</sup>. **El Pensamiento Computacional [28] abarca cinco áreas de conocimiento básicas en un contexto tecnológico: los algoritmos, los datos, las redes, la programación y el hardware.**

17 <https://ec.europa.eu/jrc/en/computational-thinking>

En el libro “Los nativos digitales no existen” [29] escribí un capítulo titulado “Eruditos digitales” que enfatiza la necesidad de enseñar tanto Pensamiento Computacional en la educación obligatoria, como de desarrollar el pensamiento crítico y habilidades de la inteligencia social, emocional y creativa que hoy en día no estamos desarrollando y que cada vez van a resultar más importantes para nuestra salud mental y nuestra coexistencia pacífica y armoniosa tanto con la tecnología como con otros humanos y con nuestro planeta. Entre otras, la habilidad para saber tomar decisiones en un entorno cambiante y con incertidumbre; la flexibilidad mental y estabilidad emocional para re-inventarse, aprender y adaptarse al cambio; la capacidad para estar aburridos, fomentando así la creatividad; la habilidad para concentrarse en una única tarea, sin interrupciones, durante períodos sostenidos en el tiempo; la capacidad para discernir entre lo que es importante de lo que no lo es, para comprobar la veracidad de las fuentes y saber formar una opinión a partir de multitud de fuentes de datos; la tolerancia hacia otros puntos de vista; y la disciplina necesaria para dotar a nuestro cuerpo biológico de las horas de descanso suficientes, sin estímulos ni interrupciones.

## La mujer y la tecnología

Asimismo, el conocimiento del ciudadano/a medio con relación a conceptos tecnológicos es muy limitado. Por ello, considero que deberíamos hacer más divulgación científico-tecnológica. Podría ser efectivo desarrollar campañas de comunicación y concienciación a través de anuncios en diversos medios de comunicación (redes sociales, periódicos, radio, TV etc.) para dar más visibilidad a la tecnología y realizar el impacto que tiene en nuestras vidas, educar en conceptos básicos e inspirar a las nuevas generaciones —sobre todo a las chicas— a estudiar carreras tecnológicas. Desgraciadamente el porcentaje de mujeres que estudian informática o que trabajan en posiciones técnicas dentro de empresas tecnológicas se encuentra entre un 10 y un 20%<sup>18</sup>. Esta falta de diversidad de género es alarmante, especialmente porque los porcentajes de mujeres en estas disciplinas ha disminuido progresivamente desde los años 80. Este declive en el porcentaje de estudiantes femeninas en informática es fruto de un conjunto de factores, entre los que destacaría:

- 1.** una imagen errónea y fuertemente estereotipada sobre quién trabaja en puestos tecnológicos y en qué consiste dicho trabajo. La imagen —reforzada por las películas, series de televisión y medios de comunicación— es la de un chico con gafas, con unos niveles de inteligencia social y emocional muy bajos, con poca higiene personal, trabajando en un sótano sin ventanas, rodeado de “comida basura”, tecleando delante de una o varias pantallas de ordenador. Esta imagen es, sin duda, muy poco atractiva para las chicas y además lejana de la realidad; y una fuerte estereotipación de género en los juguetes, libros, ropa y películas/series que consumen nuestros niños y adolescentes;
- 2.** los sesgos —conscientes y subconscientes— de género que tanto hombres como mujeres tenemos y que conllevan una infravaloración de las mujeres frente a homólogos masculinos con cualificaciones idénticas [30];
- 3.** una falta sistemática de reconocimiento hacia las mujeres en todos los contextos y en particular en las carreras tecnológicas, ilustrada por la brecha salarial o la pobre presencia de las mujeres en los premios, distinciones y posiciones de poder en este campo —por ejemplo, **el Premio Turing (equivalente al Premio Nobel en Informática)**

18 <https://smallbiztrends.com/2018/03/women-in-technology-statistics.html>

**ha sido otorgado únicamente a 3 mujeres en contraposición a 62 hombres desde su creación en 1966. Además, las mujeres fundadoras de startups solo recibieron un 2% de las inversiones de capital riesgo de EEUU en 2017<sup>19</sup>** y los equipos femeninos fundadores de startups atraen en media 82\$ de inversión por cada 100\$ de inversión obtenidos por equipos masculinos<sup>20</sup>, a pesar de que las startups fundadas por mujeres tienen mejores resultados que aquellas fundadas por hombres<sup>21</sup>;

- 4. la inexistencia de referentes femeninos en este campo que puedan inspirar a niñas y adolescentes a estudiar estas carreras; y**
- 5. una cultura marcadamente sexista y misógina, llamada **cultura brogrammer**<sup>22</sup>**, que en un estudio reciente publicado por la revista Fortune<sup>23</sup> fue el segundo motivo más frecuente –citado por un 68% de las mujeres— causante del abandono de su trabajo en el sector tecnológico (después de la maternidad).

Para fomentar la presencia de mujeres en el este campo han surgido —tanto a nivel nacional como internacional— decenas de iniciativas de apoyo, capacitación y sensibilización. En esta casa, tenemos la iniciativa de **Mujer e Ingeniería**, puesta en marcha por esta Real Academia y liderada por Sara Gómez, cuyo objetivo es, entre otros, dar apoyo a las chicas que están estudiando ingeniería a través de redes de tutorización. Otros ejemplos nacionales incluyen la iniciativa Wisibizalas de la UPF, concurso dirigido a colegios/institutos para generar reflexión sobre la visibilidad de las mujeres, dar visibilidad a mujeres que trabajan en tecnología y potenciar las referencias femeninas; la Asociación de Mujeres Investigadoras y Tecnólogas (AMIT), cuyo objetivo es promover la equidad de género; **MujeresTech**, una asociación sin ánimo de lucro que ofrece recursos y conocimiento con el objetivo de aumentar la presencia femenina en el sector digital; las **Top100, una iniciativa**

19 <http://fortune.com/2018/01/31/female-founders-venture-capital-2017/>

20 <https://techcrunch.com/2018/01/15/the-portion-of-vc-backed-startups-founded-by-women-stays-stubbornly-stagnant/?guccounter=1>

21 <https://qz.com/work/1313962/why-women-run-startups-produce-more-revenue/>

22 <https://en.wikipedia.org/wiki/Brogrammer>

23 <http://fortune.com/2014/10/02/women-leave-tech-culture/>

**dedicada a identificar las 10 mujeres más influyentes en España en 10 categorías**, que aunque no tiene una categoría tecnológica por se, incluye una categoría de emprendedoras y otra de académicas/investigadoras; o el **Premio Ada Byron** a la Mujer Tecnóloga de la Universidad de Deusto (que tuve el honor de recibir en 2016), cuyo objetivo es premiar y visibilizar trayectorias excelentes de mujeres en diversos campos tecnológicos.

Estas iniciativas son necesarias para fomentar una mayor diversidad en la tecnología. Sin embargo, quizás sea necesario aunar esfuerzos para que las iniciativas tengan un mayor impacto y visibilidad.

**La introducción de Pensamiento Computacional como asignatura troncal contribuiría asimismo a reducir la brecha de género en las carreras tecnológicas ya que todos los estudiantes (chicos y chicas por igual) adquirirían con naturalidad competencias tecnológicas, derribando así el estereotipo de género asociado a la tecnología.**

## **Empoderamiento: compartiendo**

Otra dimensión para el empoderamiento social a través de la Inteligencia Artificial conlleva la generación de espacios de colaboración (laboratorios urbanos o **living labs**) basados en datos (**compartiendo datos**), tecnología (**compartiendo infraestructuras**) y capacidades (**compartiendo talento**) para contribuir al progreso de la Inteligencia Artificial y democratizar su acceso, equilibrando desigualdades y la asimetría a la que hice referencia anteriormente. Algunos ejemplos incluyen el Mobile Territorial Lab [31] en Trento —con quienes colaboré cuando era Directora Científica en Telefónica— y el Laboratorio Urbano de Bogotá —impulsado por la ONG Data-Pop Alliance, donde soy investigadora jefa de datos. En el Mobile Territorial Lab lideré un proyecto cuyo objetivo era entender, desde un punto de vista centrado en las personas, el valor monetario que los usuarios asignan a los datos captados por sus móviles. Este proyecto recibió premio al mejor artículo científico en el

congreso internacional ACM Ubicomp 2014 [32], así como un gran interés por parte de los medios de comunicación nacionales e internacionales, al ser el primer trabajo de investigación que exploraba esta cuestión.

**La desigualdad social viene de la mano de la existencia de la propiedad.** Cuando dicha propiedad no está equitativamente distribuida en la sociedad, surge la desigualdad. El siglo XXI se caracteriza por un crecimiento de la desigualdad. Hoy en día, según un estudio reciente de Credit Suisse<sup>24</sup>, **el 1% más rico del planeta posee la mitad de la riqueza mundial y las 100 personas más ricas del mundo tienen más que las 4 mil millones de personas más pobres del mundo.** Esta acumulación de la riqueza en las manos de muy pocos ha sido atribuida parcialmente al desarrollo tecnológico y la Cuarta Revolución Industrial. Con la Revolución Agraria en el neolítico y durante miles de años, la propiedad de la tierra conllevaba riqueza. En la Revolución Industrial, la riqueza pasó a estar ligada a la propiedad de las fábricas y las máquinas. Hoy en día, podríamos argumentar que los datos y la capacidad para sacar partido de estos son el activo que más riqueza genera, dando lugar a lo que se conoce como la “**economía de los datos**”. Por tanto, si queremos maximizar el impacto positivo de la IA en la sociedad, dado que dicha inteligencia necesita de datos para poder aprender, deberíamos plantearnos nuevos modelos de propiedad, gestión y regulación de los datos. La Regulación General Europea para la Protección de Datos (RGDP) es un ejemplo en esta dirección. **Sin embargo, la complejidad en su implementación y aplicación práctica pone de manifiesto la dificultad para definir e implementar el concepto de “propiedad” cuando hablamos de un bien intangible, distribuido, variado, creciente, dinámico y replicable infinitas veces a coste prácticamente cero.**

A pesar de las dificultades, en los últimos 3-5 años han surgido distintas iniciativas a nivel mundial para promover el Big Data y la Inteligencia Artificial para el bien común, una de mis áreas de investigación desde hace aproximadamente 10 años, y así contribuir a la democratización en el acceso y uso de los datos, incluyendo:

—**New Deal on Data**, liderado por el profesor Sandy Pentland —mi director

24 <https://www.credit-suisse.com/corporate/en/research/research-institute/global-wealth-report.html>

de tesis— desde el Foro Económico Mundial, enfocado en consensuar políticas e iniciativas para que la ciudadanía tenga control sobre la posesión, el uso y la distribución de sus datos personales.

—**Flowminder**<sup>25</sup> es una ONG basada en Suecia —con quienes colabro desde hace años— con una década de experiencia en proyectos de análisis de Big Data para el bien social, con casos de éxito en el uso de Big Data para, por ejemplo, cuantificar los desplazamientos y asentamientos humanos tras los terremotos en Haití y Nepal.

—**Naciones Unidas: el World Data Forum y el Global Partnership for Sustainable Development Data.** Esta última es una alianza, liderada por Naciones Unidas, para la consecución de los 17 Objetivos de Desarrollo Sostenible (ODS) a través del análisis de datos, con más de 150 colaboradores que representan un amplio espectro de productores y usuarios de datos, incluidos gobiernos, empresas, universidades, ONG, grupos de la sociedad civil, fundaciones, oficinas de estadísticas, etc. Los datos y la Inteligencia Artificial pueden utilizarse tanto para analizar el cumplimiento de los 17 ODS como para contribuir a conseguirlos y mejorar la toma de decisiones, por ejemplo, en las políticas públicas. La colaboración entre los diferentes organismos implicados es clave para el desarrollo de las metas establecidas. Dicha colaboración solo será posible si todas las entidades partícipes (gobiernos, organismos internacionales, empresas, universidades, sociedad civil, etc.) se comprometen a trabajar conjuntamente aprovechando el potencial que brindan los datos e involucrando a los ciudadanos. En Naciones Unidas también existe desde 2009 una unidad dedicada al análisis de datos usando técnicas de Inteligencia Artificial para el bien social llamada United Nations Global Pulse<sup>26</sup>, con quienes también he colaborado. En 2014 llevamos a cabo un estudio donde demostramos el valor que aportan los datos agregados y anonimizados de la red de telefonía móvil, combinados con datos de imágenes vía satélite, para la detección de las zonas afectadas por inundaciones en México [33].

—**OPAL**<sup>27</sup> es un proyecto liderado por Data-Pop Alliance —donde soy investigadora jefa de datos— en colaboración con diversas entidades público-privadas, con el objetivo de aprovechar el Big Data y la Inteligencia Artificial

25 <http://flowminder.org/>

26 <https://www.unglobalpulse.org/>

27 <https://www.opalproject.org/>

para el bien social preservando la privacidad de las personas, de manera sostenible, escalable, estable y comercialmente viable. El proyecto OPAL propone que los algoritmos sean abiertos y que estos se ejecuten donde estén los datos y no al revés, democratizando el acceso a dichos datos y al conocimiento derivado de los mismos.

—**Existen numerosas iniciativas de empresas privadas cuyo objetivo es fomentar, aplicar y desarrollar proyectos que utilicen el Big Data y la Inteligencia Artificial para el bien social.** Entre otras, **LUCA-Big Data for Social Good de Telefónica** —que tuve el honor de crear e impulsar— cuyo objetivo es utilizar el Big Data para tener impacto social positivo; **BBVA Data & Analytics para el Bien Social**, por parte del BBVA; **Vodafone Big Data e Inteligencia Artificial para el Bien Social** —que he tenido el honor de crear e impulsar— con proyectos en salud pública, inclusión financiera, transporte y estadísticas oficiales en África y Europa; **Telenor Big Data for Social Good**, con proyectos de salud pública en Bangladesh y Pakistán; **Orange Data for Development**, dos retos pioneros donde Orange compartió datos agregados y anonimizados de Senegal y Costa de Marfil con cientos de equipos internacionales, enfocados en casos de uso para el bien social; y **Turkcell Data for Refugees Challenge**, en el que la operadora de telefonía Turkcell ha compartido datos agregados y anonimizados para contribuir a resolver la crisis de los refugiados.

—**Partnership on AI**<sup>28</sup> es una organización sin ánimo de lucro creada en 2016 por Microsoft, Facebook, Amazon, IBM, Google y Apple que agrupa hoy en día más de 50 organizaciones públicas y privadas, con el objetivo de estudiar y establecer mejores prácticas respecto a la Inteligencia Artificial, avanzar el conocimiento sobre Inteligencia Artificial por parte de la ciudadanía, y actuar como plataforma abierta para el debate sobre la Inteligencia Artificial y su impacto en las personas y la sociedad.

—**GSMA - Big Mobile Data for Social Good**<sup>29</sup>, liderada por el GSMA y Naciones Unidas en la que participan 20 operadoras de móvil del mundo para contribuir, a través del análisis de datos agregados y anonimizados de la red de telefonía móvil, a resolver problemas en las áreas de la salud pública y el cambio climático/desastres naturales.

28 <https://www.partnershiponai.org/>

29 <https://www.gsma.com/betterfuture/bd4sg>

—**AI for Good Global Summit de la ITU**<sup>30</sup>, cumbre internacional de las Naciones Unidas para el diálogo sobre la Inteligencia Artificial, orientada a identificar las aplicaciones prácticas de la IA para la mejora de la sostenibilidad del planeta. Está gestionada por la ITU (International Telecommunications Union) como organismo especializado de las Naciones Unidas para las Tecnologías de la Información y la Comunicación.

—**El Center for Humane Technology**<sup>31</sup> es un centro multidisciplinar recientemente creado en California que defiende el desarrollo de tecnología humana, es decir, tecnología que ha sido diseñada teniendo en cuenta los valores, las necesidades y los intereses humanos por encima de todo.

30 <https://www.itu.int/en/ITU-T/AI/2018/Pages/default.aspx>  
31 <https://humanetech.com/>

## **7. Principios éticos y dimensiones de trabajo necesarias para el desarrollo de una Inteligencia Artificial centrada en las personas**

A lo largo de la historia, el desarrollo de **tecnologías disruptivas** ha tenido impacto profundo en la vida y en las relaciones humanas. La agricultura, la imprenta, la máquina de vapor, la electricidad o internet son ejemplos de tecnologías que han afectado profundamente nuestra manera vivir, trabajar y relacionarnos.

La Inteligencia Artificial también está transformando el mundo en que vivimos. Hoy contamos con cantidades masivas de datos que podemos utilizar para entrenar algoritmos de Inteligencia Artificial que permiten a investigadores, empresas, gobiernos y otros actores del sector público abordar problemas complejos.

Decisiones con impacto tanto individual como colectivo que previamente eran tomadas por humanos —con frecuencia expertos— son tomadas hoy en día por sistemas de Inteligencia Artificial (es decir, **algoritmos**) incluyendo decisiones relativas a la contratación de personas, la concesión de créditos y préstamos, las sentencias judiciales, los tratamientos y diagnósticos médicos o la compraventa de acciones en bolsa. Las decisiones algorítmicas basadas en datos tienen el potencial de mejorar nuestra toma de decisiones. De hecho, la historia ha demostrado que las decisiones humanas no son perfectas: están sujetas a conflictos de interés, a la corrupción, al egoísmo y a sesgos cognitivos, lo que ha resultado en procesos y resultados injustos y/o inefficientes. Por tanto, el interés hacia el uso de algoritmos puede interpretarse como el resultado de una demanda de mayor objetividad en la toma de decisiones<sup>32</sup>.

Sin embargo, la toma de decisiones basadas en algoritmos entrenados con datos no está exenta de limitaciones. Curiosamente, las palabras de Platón hace 2.400 años son sorprendentemente vigentes hoy en día, “**Una buena decisión está basada en conocimiento, no en números (datos)**”.

Cuando estas decisiones afectan a miles o millones de personas, surgen **dilemas éticos importantes**<sup>33</sup>. Por ejemplo, ¿quiere esto decir que las decisiones automáticas escaparán de nuestro control? ¿Qué niveles de seguridad poseen estos sistemas para protegerse de los ciberataques o de un uso malicioso de los mismos? ¿Cómo podemos garantizar que dichas

32 <https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms>

33 La ética hace referencia al conjunto de normas y principios según los cuales determinamos cuales son los comportamientos aceptables y los que no son aceptables

decisiones y/o actuaciones no tienen consecuencias negativas para las personas? ¿Quién es responsable de dichas decisiones? ¿Qué sucederá cuando un algoritmo nos conozca a cada uno de nosotros mejor que nosotros mismos y pueda aprovechar dicho conocimiento para manipular de manera subliminal nuestro comportamiento?

Recientemente han surgido numerosos movimientos, organizaciones, nacionales o supra-nacionales —por ejemplo, en la Unión Europea<sup>34</sup>— para la definición de principios que sirvan para el establecimiento de estándares globales y regulación de este ámbito. Por ejemplo, en el congreso organizado por el Future of Life Institute en 2017, con la participación de más de 1.200 figuras internacionales relacionadas con la innovación tecnológica y científica, se definieron los Principios de Asilomar para el desarrollo de la Inteligencia Artificial con un total de 23 recomendaciones. Recientemente, la Comisión Europea ha nombrado un Comité de Expertos de Alto Nivel en Inteligencia Artificial, con un foco en las implicaciones éticas, legales y sociales de la IA<sup>35</sup>, del que soy miembro reserva. A nivel nacional, destaca la declaración de Barcelona<sup>36</sup>, impulsada por el Centro de Investigación en Inteligencia Artificial del CSIC en la que se definen 6 principios básicos para un desarrollo ético de la IA.

Mas allá del respeto a los derechos humanos fundamentales<sup>37</sup>, la literatura existente ha propuesto un conjunto de principios éticos y dimensiones de trabajo que considero es necesario abordar para conseguir que la Inteligencia Artificial tenga impacto positivo. El lector interesado puede encontrar una versión extendida de algunos de estos principios éticos y dimensiones de trabajo en un artículo reciente escrito con mis colaboradores de MIT, Data-Pop Alliance, Harvard y FBK, y publicado en la revista “Philosophy and Technology” [34]. En este capítulo resumo los más relevantes, agrupados en **5 pilares** marcados con diferentes colores [35]:

34 [http://europa.eu/rapid/press-release\\_IP-18-6689\\_en.htm](http://europa.eu/rapid/press-release_IP-18-6689_en.htm)

35 <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

36 <http://www.iiia.csic.es/barcelonadeclaration/>

37 [https://en.wikipedia.org/wiki/Fundamental\\_rights](https://en.wikipedia.org/wiki/Fundamental_rights)

## 1. Autonomía y dignidad

La autonomía es un 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. Para esto, se deben disponer de una serie de reglas que definan el comportamiento de estos sistemas de acuerdo con los principios éticos aceptados en la sociedad donde sean utilizados. Hay numerosos ejemplos de principios éticos propuestos en la literatura para este propósito<sup>38</sup>, así como institutos y centros de investigación creados con este fin, como el AI Now Institute en New York University y el Digital Ethics Lab en la Universidad de Oxford. Sin embargo, esta es un área activa de investigación y no hay un método único para la incorporación de principios éticos en los procesos algorítmicos de decisión basados en datos. Es importante destacar la importancia de que todos los 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”.

38 The ACM Code of Ethics: <https://www.acm.org/code-of-ethics>

## 2. Justicia y solidaridad

No discriminación. La justicia debería ser un elemento central en el desarrollo de sistemas de decisión (y actuación) automáticos basados en la Inteligencia Artificial. 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. Además, los procesos de decisión algorítmicos basados en datos pueden implicar que se les denieguen oportunidades a personas no por sus propias acciones, sino por las acciones de otros con los que comparten ciertas características. Por ejemplo, algunas compañías de tarjetas de crédito han reducido el límite de crédito de sus clientes no por el historial de pagos del propio cliente, sino como resultado del análisis del comportamiento de otros clientes con un historial de pagos deficiente que habían comprado en los mismos establecimientos donde el cliente había comprado. En la literatura se han propuesto diferentes soluciones para afrontar la discriminación algorítmica y maximizar la justicia<sup>39</sup>. Sin embargo, me gustaría subrayar la urgencia para que expertos y expertas de distintos campos (incluyendo el derecho, la economía, la ética, la informática, la filosofía y las ciencias políticas) inventen, evalúen y validen en el mundo real diferentes métricas de justicia algorítmica para diferentes tareas. Además de esta investigación empírica, es necesario proponer un marco de modelado teórico –avalado por la evidencia empírica—que ayude a los usuarios de dichos algoritmos a asegurarse de que las decisiones tomadas son lo más justas posible.

39 “Fairness in machine learning”. <https://fairmlclass.github.io>

## **Cooperación**

Debido a la transversalidad de la Inteligencia Artificial y su potencial aplicación a todos los ámbitos, deberíamos fomentar y desarrollar un intercambio constructivo de recursos y conocimientos entre los sectores privado, público y la sociedad para conseguir el máximo potencial de aplicación y competitividad. Esta necesidad de cooperación no solo entre diferentes sectores sino también entre naciones —dada la globalización— ha sido enfatizada por el conocido historiador y pensador israelí, el profesor Yuval Noah Harari [27].

### **3. Beneficiencia**

#### **Sostenibilidad**

El desarrollo tecnológico en general y de sistemas de Inteligencia Artificial en particular conlleva un consumo energético significativo, con impacto negativo en el medio ambiente, como he explicado anteriormente. Las técnicas de deep learning requieren elevadas capacidades de computación con costes energéticos prohibitivos, sobre todo si consideramos el despliegue de este tipo de sistemas a gran escala. Por ello, cada vez es más importante que el desarrollo tecnológico esté alineado con la responsabilidad humana de garantizar las condiciones básicas para la vida en nuestro planeta y para preservar el medio ambiente para generaciones futuras. Asimismo, la Inteligencia Artificial será clave para permitirnos abordar algunos de los retos más importantes en el contexto del medio ambiente (por ejemplo, cambio climático, la disponibilidad de recursos...) así como para permitirnos desarrollar medios de transporte (por ejemplo, coches autónomos eléctricos) y modelos energéticos más eficientes y sostenibles (por ejemplo, smart grids).

#### **Veracidad**

Hoy en día podemos crear contenido (texto, fotos, videos) sintético, inventado por algoritmos, que es indistinguible de contenido veraz. Esta capacidad ha generado la aparición del concepto de fake news o noticias falsas, que pueden definir la opinión pública respecto a temas importantes —como quien debería ser el próximo presidente del país o si debería un país pertenecer a la Unión Europea o no— para así favorecer los intereses de una minoría que tiene la habilidad de generar y disseminar dichas noticias falsas. Por ello, el principio de la veracidad tanto de los datos usados para entrenar algoritmos de IA como de los contenidos que consumimos es de suma importancia.

## Diversidad

Dada la variedad de casos de uso en los que puede aplicarse la Inteligencia Artificial es importante reflexionar sobre la frecuente falta de diversidad en los equipos que crean los sistemas de Inteligencia Artificial, hasta ahora compuestos mayoritariamente por grupos homogéneos de profesionales de la informática. En el futuro, deberíamos asegurarnos de que los equipos son diversos tanto con relación a las áreas de conocimiento como con relación a factores demográficos —y en particular de género, donde el porcentaje de mujeres profesionales en la informática es menor de un 20% en la mayoría de las empresas tecnológicas, como he explicado anteriormente.

Asimismo, los algoritmos de personalización y recomendación adolecen con frecuencia de falta de diversidad en sus resultados, y tienden a encasillar a sus usuarios en ciertos patrones de gustos lo que da lugar a la aparición de lo que Pariser ha denominado el filtro burbuja [38]. Esta falta de diversidad en la personalización/recomendación no es deseable, ya que limita las oportunidades de la tecnología para ayudarnos a descubrir contenido (por ejemplo, películas, libros, música, noticias o amigos) diferente a nuestros propios gustos y por tanto contenido que nos ayuda a entender otros puntos de vista y fomentar la apertura de mente.

## 4. Explicabilidad

### Transparencia

La transparencia hace referencia a la cualidad de poder entender un modelo computacional y por ello puede ser un mecanismo que contribuya a la atribución de responsabilidad de las consecuencias del uso de dicho modelo. Un modelo es transparente si una persona puede observarlo y entenderlo con facilidad. Burrell [35] propone tres tipos distintos de opacidad —i.e. falta de transparencia— en las decisiones algorítmicas: (1) opacidad intencional, cuyo objetivo es la protección de la propiedad intelectual de los inventores de los algoritmos. Este tipo de opacidad podría mitigarse con legislación que obligaría al uso de sistemas de software abierto. La nueva Regulación General Europea de Protección de Datos (RGPD)<sup>40</sup> con su derecho a una explicación es un ejemplo de este tipo de legislación. Sin embargo, intereses comerciales y gubernamentales poderosos pueden dificultar la eliminación de este tipo de opacidad; (2) opacidad de conocimiento, debido al hecho de que la gran mayoría de las personas carecen de las habilidades técnicas para entender cómo funcionan los algoritmos y modelos computacionales construidos a partir de los datos. Este tipo de opacidad se vería atenuada con programas educativos en competencias digitales —como he explicado anteriormente— y permitiendo que expertos independientes pudiesen aconsejar a aquellos afectados por los procesos de decisión algorítmicos basados en datos; y (3) opacidad intrínseca, que surge por la naturaleza de ciertos métodos de aprendizaje por ordenador (por ejemplo, modelos de deep learning o aprendizaje profundo). Esta opacidad es bien conocida en la comunidad de investigación de aprendizaje computacional y también se conoce como el problema de la interpretabilidad.

40 <https://www.eugdpr.org>

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 —lo cual queda contemplado en la RGPD a nivel europeo— sino también respecto en qué situaciones los humanos están interaccionando con sistemas artificiales (por ejemplo, chatbots) vs con otros humanos.

### **Responsabilidad y el papel del humano**

También es importante tener claridad con relación a la atribución de responsabilidad de las consecuencias de las acciones o decisiones de sistemas autónomos, del mismo modo que sucede con el resto de los productos utilizados en la sociedad. La transparencia suele considerarse un factor fundamental para contribuir a la claridad de atribución de responsabilidad. Sin embargo, la transparencia y las auditorías no son suficientes para garantizar una clara responsabilidad. De hecho, en un artículo reciente, Kroll [36] ha propuesto el uso de métodos computacionales para proporcionar claridad con relación a la atribución de responsabilidad, incluso cuando parte de la información está oculta.

Finalmente, creo que es constructivo tener una visión sinérgica entre la Inteligencia Artificial y el ser humano. Esta visión suele llamarse “**aumento de la inteligencia**” (intelligence augmentation) de manera que los sistemas de Inteligencia Artificial se utilizan para aumentar o complementar la inteligencia humana. Por ejemplo, un buscador de Internet puede considerarse un sistema de aumento de nuestra inteligencia ya que expande nuestro conocimiento con la capacidad de procesar miles de millones de documentos y encontrar los más relevantes; o un sistema de traducción simultánea automática, ya que permite a personas que no hablan el mismo idioma poderse comunicar.

## 5. No maleficencia

### **Fiabilidad y seguridad**

La gran mayoría —si no todos— los sistemas, productos y bienes que utilizamos (por ejemplo, alimentos, electrodomésticos, vehículos, ropa, juguetes, medicinas, aparatos médicos, maquinaria industrial...) están sujetos a estrictos controles de calidad, seguridad y fiabilidad para minimizar el potencial impacto negativo que puedan tener en la sociedad. Del mismo modo, es de esperar que los sistemas de Inteligencia Artificial también estén sujetos a procesos similares. Mas allá de los procesos teóricos de seguridad, verificación y fiabilidad, quizás tendría sentido crear una autoridad que a nivel europeo certificase la calidad, seguridad y fiabilidad de los sistemas de Inteligencia Artificial antes de que fuesen comercializados o implementados en la sociedad. Asimismo, los sistemas autónomos deberían velar por la seguridad e integridad tanto de las personas que los utilizan o se ven afectadas por su acción como por su propia seguridad frente a la manipulación y los ciberataques.

### **Reproducibilidad**

Para generar confianza, los sistemas deberían tener consistencia en su operación de modo que su comportamiento sea no solo entendible por parte de un humano, sino también reproducible, es decir, replicable cuando es sometido a los mismos datos de entrada o a la misma situación/contexto.

## **Prudencia**

La aplicación de la Inteligencia Artificial exige que los profesionales cumplan requisitos estrictos para su desarrollo, tales como garantizar la disponibilidad de datos suficientes (e idealmente de calidad), el análisis de las hipótesis de trabajo desde diversas perspectivas y la disponibilidad de expertos/as y recursos para poder hacer un análisis e interpretación de los modelos y sus resultados. El principio de la prudencia destaca la importancia de considerar en las fases iniciales de diseño de cualquier sistema las diversas alternativas y opciones existentes para maximizar su impacto positivo y minimizar los potenciales riesgos y consecuencias negativas derivadas de su aplicación.

## **Protección de datos y privacidad**

En un mundo de datos, en el que generamos y consumimos datos de forma ubicua y masiva, los derechos a la protección de los datos personales y al respeto de la privacidad son cuestionados e incluso llevados al límite constantemente. Numerosos estudios se han enfocado en el mal uso de datos personales proporcionados por los usuarios de servicios y en la agregación de datos de diferentes fuentes por parte de entidades como los data brokers con implicaciones directas en la privacidad de las personas. Un elemento que a menudo no se tiene en cuenta es que los avances en los algoritmos de aprendizaje automático, combinados con la disponibilidad de nuevas fuentes de datos sobre el comportamiento humano (por ejemplo, datos de social media), permiten la inferencia de información privada (por ejemplo, orientación sexual, inclinaciones políticas, nivel de educación o niveles de estabilidad emocional) que nunca ha sido explícitamente revelada por las personas. En un proyecto de investigación reciente demostramos que a partir de datos no personales éramos capaces de inferir atributos tan personales como algunas dimensiones de la personalidad, el nivel de educación o los intereses [37]. Este elemento es esencial para entender las implicaciones del uso de algoritmos

para modelar —o incluso influenciar— el comportamiento humano a nivel individual, como ha quedado patente en el escándalo reciente de Facebook/Cambridge Analytica<sup>41</sup>. Por ello, considero que debería haber atributos y características de las personas que tendrían que permanecer en la esfera privada (por ejemplo, orientación sexual, religión...) —a no ser que la persona decida que sea así expresamente— y no deberían utilizarse o inferirse en los sistemas de IA.

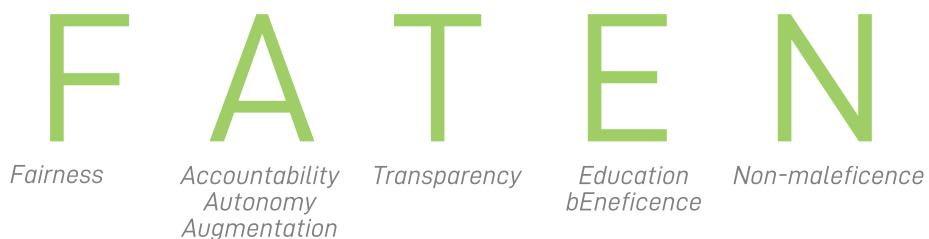
Europa ha asumido cierto liderazgo a nivel mundial con la reciente entrada en vigor del Reglamento General de Protección de Datos o RGPD, que se suma a derechos fundamentales como los derechos a establecer y desarrollar relaciones con otros seres humanos, a la desconexión tecnológica y a estar libre 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.

41 [https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge\\_Analytica\\_data\\_scandal](https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal)

Podemos captar los principios anteriores en el acrónimo inglés FATEN, haciendo referencia a: la justicia (F de fairness, solidaridad y cooperación); la atribución de responsabilidad, asimetría y aumento de la inteligencia (A de accountability y de asymmetry); la transparencia (T de transparency); la bEneficencia y la educación, contribuyendo al progreso con sostenibilidad y diversidad (E de bEneficece y de education); y la no-maleficencia (N de non-maleficence), preservando la privacidad, con seguridad, veracidad, fiabilidad y prudencia. Muchos de estos principios han sido descritos en la literatura, como en [39].

En suma, deberíamos siempre centrar el desarrollo de los sistemas de IA en las personas y fomentar la creación de entornos colaborativos para experimentar y co-crear políticas y soluciones basadas en la IA, consensuadas por los humanos.

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





Desconozco qué nos deparará el futuro y dónde estaremos dentro de veinte años. Pero sí puedo soñar cómo me gustaría que fuese. Por ello, concluyo mi discurso compartiendo con Vds. tres sueños para el futuro.

- 1. En primer lugar, deseo que sea un futuro donde la tecnología en general —y la Inteligencia Artificial en particular— forme parte integral de nuestras vidas, donde co-existamos sinérgica y armónicamente con tecnología que nos ayude no solo a vivir más, sino sobre todo a vivir mejor, a todos. El potencial para tener impacto positivo es inmenso y considero que no deberíamos desaprovechar esta oportunidad. Sin embargo, no es un futuro cierto tal y como he articulado anteriormente. Por ello debemos afrontar seriamente los retos y las limitaciones que presentan los sistemas de Inteligencia Artificial actuales —anteriormente descritos— para conseguir que este sueño se haga realidad. Espero no solo poder formar parte de ese futuro sino también y muy especialmente poder contribuir al mismo con mi trabajo.**
- 2. En segundo lugar, sueño con que España invierta mucho más en Inteligencia Artificial que lo que hace actualmente, cumpliendo los principios explicados anteriormente, para convertirse en líder en Europa y puente con Latinoamérica y África. Idealmente, destacaríamos por una apuesta ambiciosa para conseguir la adopción de la IA en nuestras empresas y administraciones públicas, con un refuerzo notable de nuestra excelencia científica en esta área, invirtiendo significativamente en la formación, atracción y retención del talento, en la actualización de nuestro sistema educativo y en el fomento de la creatividad y la innovación. Tenemos la oportunidad de elevar no solo el crecimiento empresarial y económico, sino sobre todo nuestra calidad de vida. Espero y deseo que no dejemos escapar esta oportunidad.**
- 3. Finalmente, espero, deseo —y sueño— que cada vez haya más mujeres ingenieras, investigadoras, inventoras, innovadoras en tecnología y académicas de esta Real Academia, que sean excepcionales no por su condición de ser mujeres, sino por la brillantez de sus ideas y el impacto de su trabajo.**

Gracias por contribuir con este reconocimiento a que al menos uno de mis sueños se haya hecho realidad.

## **9. Introducción al discurso de recepción de condición de Académica en la Real Academia de Ingeniería**

**Excelentísimo Sr Presidente de la Real Academia de Ingeniería,  
Excelentísimas Señoras y Señores académicos,  
Señoras y Señores, queridos amigos y amigas:**

No puedo comenzar este discurso sin agradecer sincera y profundamente a todos los miembros de esta Real Academia su apoyo a mi candidatura. En particular, agradezco efusivamente a los académicos Javier Aracil, Josefina Gomez, Manuel Marquez y Elias Muñoz que promovieron mi candidatura y también a aquellos que la avalaron con su voto. Comencé mi relación con la Real Academia de Ingeniería con anterioridad a ser propuesta. Por ello, me gustaría destacar el apoyo recibido por Javier Pérez de Vargas y Sara Gómez en las primeras actividades en las que he participado en esta casa. Finalmente, agradezco enormemente al presidente, Elías Fereres, por guiarme en mis primeros pasos, por su confianza, por su valiosa y constructiva ayuda con este discurso y por su gran dosis de paciencia conmigo.

Gracias por el interés hacia mi perfil y por la valentía por promover y apoyar a una candidatura que es cuando menos original: soy mujer (la cuarta de sesenta académicos), la primera ingeniera de telecomunicación y quizás de insuficiente edad.

Además, me siento especialmente honrada por recibir hoy una medalla tan memorable como la nº1 de esta Real Academia, de la que fue depositario el ilustre académico constituyente D. Eugenio Andrés Puente, a quien la Academia rindió un homenaje póstumo el pasado 25 de septiembre. El profesor Andrés Puente fue una figura señera de la Escuela Técnica Superior de Ingenieros Industriales de la UPM y un pionero de la investigación en ingeniería industrial en España. Además de su excelente trayectoria docente y su nutrida producción científica, destacó por su dedicación al fomento de la innovación industrial y su estrecha colaboración con la industria, como también es mi caso. Por ello, recibió galardones tan prestigiosos como la Encomienda con placa de la Orden Civil de Alfonso X el Sabio y la Cruz de primera clase de la Orden de Mérito de la República Federal Alemana, donde obtuvo su doctorado. Así pues, recibo esta medalla, que fue la suya, con mucha ilusión, con humildad, con un gran sentido de la responsabilidad y con el firme compromiso de ser digna de tomar su testigo.

Me gustaría compartir esta distinción que hoy se me otorga con todas las personas e instituciones que siempre me han apoyado y han creído en mí. Comenzando por mis padres —por su amor, generosidad y sabiduría sin límites, por haberme dado alas para volar y al mismo tiempo haberme ofrecido siempre una sólida rama a la que regresar—mis hermanos y el resto de familiares, mis amigos y amigas, maestras y profesores, colegas, jefes y mentores, los investigadores e investigadoras (pasados y presentes) con los que he colaborado durante mi carrera científica y, como no, mis tres hijos maravillosos (Matías, Adrian y Sofía) y mi marido, Kristof, “my partner in crime”, como decimos en inglés, compañero incansable en esta aventura intensa y llena de sorpresas —como es este nombramiento—que es la vida...como dijo Mark Twain, “para obtener el valor completo de una alegría, tienes que tener alguien con quien compartirla”. En mi caso, sin duda, las alegrías son plenas.

Eleanor Roosevelt dijo que el futuro pertenece a aquellos que creen en la belleza de sus sueños. Y coincido con las palabras de Anaïs Nin: los sueños son necesarios para la vida. De hecho, en mi vida, no han faltado ni faltan los sueños, sueños que me han inspirado, me han animado y me han ayudado a encontrar el camino. Por ello, hoy me gustaría compartir con Vds. algunos de mis sueños profesionales, tanto aquellos que con el tiempo se han convertido, para mi sorpresa, en realidad, como otros que me encantaría hacer posibles en un futuro.

Meses antes de que fuese creada la Real Academia de Ingeniería, cuando yo era estudiante de Teleco en la UPM, descubrí la Inteligencia Artificial, el área a la que he dedicado mi vida profesional. Y fue amor a primera vista. Comencé utilizando estaciones de trabajo de SGI que tenían menos capacidad de computación que muchos de los teléfonos móviles de hoy. La tecnología, con su velocidad de progreso exponencial, nos hace sentir aún más viejos de lo que somos.

En aquel momento cobró fuerza y forma mi primer gran sueño: marcharme a EE.UU. a cursar estudios de doctorado en Inteligencia Artificial.

Tuve la suerte de poder realizar este sueño, gracias a una beca de la Fundación Obra Social la Caixa y a la oportunidad que me brindó mi director de tesis, el profesor Sandy Pentland de MIT, para hacer un doctorado en el mítico Media Lab. Con Sandy comencé mi carrera científica, publiqué mis primeros artículos científicos e impartí mis primeras charlas. Siempre estaré agradecida a Sandy

por la oportunidad, por su valioso apoyo, enseñanzas y consejo durante estas más de dos décadas. Tras el doctorado comencé mi carrera profesional como investigadora en Microsoft Research. Aunque siempre deseé poder regresar a España, durante mis más de doce años en EE. UU. nunca pensé que algún día podría hacerlo. Por ello, durante mucho tiempo un posible regreso no fue más que otro de mis sueños.

Sin embargo, hace once años, la Fortuna —bueno, más bien Telefónica— de nuevo llamó a mi puerta con una oportunidad para ver cumplido ese sueño: regresar como Directora Científica —la primera mujer— en Telefónica I+D en Barcelona.

Dicen que la vida es cíclica (el círculo de la vida) y sin duda en mi caso así ha sido, ofreciéndome la oportunidad de cumplir otro sueño que consideré que nunca vería cumplido: regresar a Alicante, mi tierra. Sin embargo, hace 3 años, por motivos personales, nos trasladamos a vivir a Alicante para estar junto a mi familia. Gracias a la tecnología y al apoyo de las organizaciones en las que he trabajado y trabajo —y muy especialmente Vodafone— he podido convertir en realidad lo que durante décadas parecía imposible.

Reitero mi agradecimiento a todos mis colaboradores y mentores durante mi carrera profesional, y en especial al Prof. Sandy Pentland, a Eric Horvitz, Mary Czerwinski, Jennifer Chayes, Carlos Domingo, Emmanuel Letouzé y Katia Walsh, por su valioso consejo y la confianza plena que siempre han depositado en mí.

Este reconocimiento es para todos ellos, porque sin su ayuda, apoyo y colaboración no habría sido capaz de conseguir lo que hemos conseguido juntos.

Y como no, agradezco a todos los presentes el esfuerzo que han hecho para poder estar hoy aquí compartiendo este momento tan solemne y memorable conmigo.

Desde aproximadamente 25 años, he centrado mi actividad investigadora en el modelado computacional del comportamiento humano utilizando técnicas de inteligencia artificial. He trabajado con habitaciones, ropas, coches y móviles inteligentes. He inventado sistemas capaces de reconocer comportamientos o características humanas, tales como: expresiones faciales [1], actividades [2], interacciones humanas [3], maniobras en la conducción [4], apnea del sueño [5], riesgo crediticio [6], puntos calientes de crimen en las ciudades [7] o

incluso el aburrimiento [8]...he construido sistemas interactivos e inteligentes en ordenadores y en móviles, por ejemplo [9] [10]. Gracias a haber desarrollado mi carrera científica en un contexto empresarial, he podido sentir, en primera persona, la felicidad profunda que te inunda cuando consigues que lo que no era más que una idea —a veces incluso un tanto alocada— se convierta en una realidad que puede ayudar a millones de personas.

He sido no solo testigo sino también partícipe del progreso tecnológico, de la presencia cada vez más relevante y ubicua de la tecnología en nuestras vidas y de la co-dependencia que hemos desarrollado con la misma.

Durante toda mi carrera, la inspiración y el motor de mi trabajo ha sido siempre una pregunta con una clara aplicación social. En todos mis proyectos, la persona —a nivel individual o colectivo— ha sido y es el elemento central: tecnología por y para la sociedad; tecnología que nos entiende, como paso previo para poder ser capaz de ayudarnos. Sin embargo, hoy, más que nunca, siento preocupación por el impacto que dicha tecnología está teniendo en nuestras vidas, impacto no necesariamente siempre positivo<sup>1</sup>.

Por tanto, además de compartir algunos de mis sueños profesionales —pasados y presentes— he decidido dedicar mi discurso de recepción de condición de Académica en la Real Academia de la Ingeniería a la Inteligencia Artificial, haciendo un breve recorrido por su historia y aportando mi visión del futuro.

1 <http://humanetech.com/>





- [1] N. Oliver, A. Pentland and F. Bérard, "LAFTER: a real-time face and lips tracker with facial expression recognition," *Pattern Recognition*, vol. 33, no. 8, pp. 1369-1382, 2000.
- [2] N. Oliver, E. Horvitz and A. Garg, "Layered representations for human activity recognition," in *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*, Pittsburgh, PA, USA, 2002.
- [3] N. Oliver, B. Rosario and S. Pentland, "A Bayesian computer vision system for modeling human interactions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 831-843, 2000.
- [4] N. Oliver and A. Pentland, "Graphical models for driver behavior recognition in a SmartCar," in *Proceedings of the IEEE Intelligent Vehicles Symposium 2000*, Dearborn, MI, USA, 2000.
- [5] N. Oliver and F. Flores-Mangas, "HealthGear: a real-time wearable system for monitoring and analyzing physiological signals," in *International Workshop on Wearable and Implantable Body Sensor Networks*, Cambridge, MA, USA, 2006.
- [6] J. San Pedro, D. Proserpio and N. Oliver, "MobiScore: towards universal credit scoring from mobile phone data," in *International Conference on User Modeling, Adaptation, and Personalization*, 2015.
- [7] A. Bogomolov, B. Lepri, J. Staiano, N. Oliver, Pianesi and A. F. and Pentland, "Once upon a crime: towards crime prediction from demographics and mobile data," in *ACM Int Conf on Multimodal Interaction (ICMI)*, 2014.
- [8] M. Pielot, T. Dingler, J. San Pedro and N. Oliver, "When attention is not scarce-detecting boredom from mobile phone usage," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Osaka, Japan, 2015.
- [9] R. de Oliveira, M. Cherubini and N. Oliver, "MoviPill: improving medication compliance for elders using a mobile persuasive social game," in *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Copenhagen, Denmark, 2010.
- [10] N. Oliver and F. Flores-Mangas, "MPTrain: a mobile, music and physiology-based personal trainer," in *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, Helsinki, Finland, 2006.
- [11] D. Rumelhart, G. E. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533-536, 1986.
- [12] N. Oliver, A. Pentland and F. Berard, "LAFTER: lips and face real time tracker," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, USA , 1997.
- [13] M. Brand, N. Oliver and S. Pentland, "Coupled hidden Markov models for complex action recognition," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, USA, 1997.
- [14] Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436-444, 2015.
- [15] A. Karatzoglou, X. Amatriain, L. Baltrunas and N. Oliver, "Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering," in *Proceedings of the fourth ACM conference on Recommender systems*, Barcelona, Spain, 2010.

- [16] Y. Shi, A. Karatzoglou, L. Baltrunas, M. Larson, N. Oliver and A. Hanjalic, "CLiMF: learning to maximize reciprocal rank with collaborative less-is-more filtering," in *ACM conference on Recommender systems*, New York, 2012.
- [17] X. Amatriain, N. Lathia, J. Pujol, H. Kwak and N. Oliver, "The wisdom of the few: a collaborative filtering approach based on expert opinions from the web," in *32nd international ACM SIGIR conference on Research and Development in Information Retrieval*, 2009.
- [18] M. Vieira, V. Frias-Martinez, N. Oliver and E. Frias-Martinez, "Characterizing Dense Urban Areas from Mobile Phone-Call Data: Discovery and Social Dynamics," in *IEEE Second International Conference on Social Computing*, Minneapolis, MN, USA, 2010.
- [19] J. Froelich, J. Neumann and N. Oliver, "Sensing and Predicting the Pulse of the City through Shared Bicycling," in *Proceedings of Twenty-First International Joint Conference on Artificial Intelligence*, 2009.
- [20] P. Jawandhiya, "Hardware design for machine learning," *International Journal of Artificial Intelligence and Applications (IJAIA)*, vol. 9, no. 1, pp. 63-84, 2018.
- [21] A. Andrae, "Total Consumer Power Consumption Forecast," in *Nordic Digital Business Summit*, 2017.
- [22] N. Bostrom, *SuperIntelligence: Paths, dangers, strategies*, Oxford, UK: Oxford University Press, 2014.
- [23] J. Hawkins, "What intelligent machines need to learn from the neocortex," *IEEE Spectrum*, 2017.
- [24] S. Ahmad and J. Hawkins, "Properties of Sparse Distributed Representations and their Application to Hierarchical Temporary Memory," Arxiv, 2015.
- [25] D. Lazer, M. Baum, Y. Benkler, A. J. Berinsky, K. Greenhill, F. Menczer and e. al, "The science of fake news," *Science*, pp. 1094-1096, 2018.
- [26] M. Arntz, T. Gregory and U. Zierahn, "The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis," *OECD Social, Employment and Migration Working Papers*, vol. 189, no. OECD Publishing , 2016.
- [27] Y. N. Harari, *21 lessons for the 21st century*, London: Penguin Random House, UK, 2018.
- [28] S. Bocconi, S. Chioccariello, G. Dettori, A. Ferrari and K. Engelhardt, "Developing computational thinking in compulsory education," *JCR Science for Policy Report*, 2016.
- [29] S. Lluna and J. Pedreira, *Los nativos digitales no existen*, Deusto Editorial, 2017.
- [30] R. Steinpreis, K. Anders and D. Ritzke, "The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study," *Sex Roles*, vol. 41, no. 7--8, pp. 509-528, 1999.
- [31] S. Centellegher, M. De Nadai, M. Caraviello, C. Leonardi, M. Vescovi, Y. Ramadian, N. Oliver, F. Pianesi, A. Pentland, F. Antonelli and B. Lepri, *The Mobile Territorial Lab: a multilayered and dynamic view on parents' daily lives*, EPJ Data Science, 2016.

- [32] J. Staiano, N. Oliver, B. Lepri, R. de Oliveira, M. Caraviello and N. Sebe, "Money walks: a human-centric study on the economics of personal mobile data," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Seattle, WA, USA , 2014.
- [33] Y. Torres Fernández, D. Pastor Escuredo, A. Morales Guzmán, J. Baue, A. Wadhw, C. Castro Correa, L. Romanoff, J. Lee, A. Rutherford, V. Frias Martínez, N. Oliver, E. Frias-Martinez and M. Luengo-Oroz, "Flooding through the lens of mobile phone activity," in *Proceedings of IEEE Global Humanitarian Technology Conference*, San Jose, CA, USA , 2014.
- [34] B. Lepri, N. Oliver, E. Letouzé, A. Pentland and P. Vinck, "Fair, Transparent, and Accountable Algorithmic Decision-making Processes," *Philosophy & Technology*, pp. 1-17, 2017.
- [35] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data and Society*, 2016.
- [36] J. Kroll, "Accountable Algorithms," PhD Dissertation in the Computer Science Department of Princeton, Princeton, US, 2015.
- [37] S. Park, A. Matic, K. Garg and N. Oliver, "When Simpler Data Does Not Imply Less Information: A Study of User Profiling Scenarios With Constrained View of Mobile HTTP [S] Traffic," *ACM Transactions on the Web (TWEB)*, vol. 12, no. 9, 2018.
- [38] E. Pariser, *The filter bubble: how the personalized web is changing what we read and how we think*, New York: Penguin Books, 2012.
- [39] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Shafer, P. Valcke and E. Vayena, "An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles and Recommendations," *Minds and Machines*, no. December, 2018.



Nuria Oliver es Directora de Investigación en Ciencias de Datos en Vodafone, Chief Data Scientist en Data-Pop Alliance y Chief Scientific Advisor para el Vodafone Institute. Es ingeniera superior de Telecomunicación por la UPM y doctora por el Massachusetts Institute of Technology (MIT) en Inteligencia Perceptual. Tiene más de 20 años de experiencia investigadora en MIT, Microsoft Research (Redmond, WA) y como primera directora científica (mujer) en Telefónica I+D y en Vodafone.

Su trabajo en el modelado computacional del comportamiento humano, la interacción persona-máquina, la informática móvil y el análisis de Big Data —especialmente para el Bien Social— es internacionalmente conocido con más de 180 publicaciones científicas, citadas más de 15.000 veces y con una decena de premios y nominaciones a mejor artículo científico.

Nuria es la única investigadora española reconocida por el ACM como Distinguished Scientist, Fellow y miembro de CHI Academy. También es Fellow del IEEE y de la Asociación Europea de Inteligencia Artificial. Es doctora Honoris Causa por la Universidad Miguel Hernández y académica de la Real Academia de Ingeniería y de la Academia Europaea.

Nuria es miembro del comité científico asesor de seis universidades europeas. Además, asesora al Gobierno de España y a la Comisión Europea en temas relacionados con la Inteligencia Artificial, a Mahindra Comviva y al Future Digital Society. Es miembro de un Global Future Council del Foro Económico Mundial.

Es miembro del comité de programa de los principales congresos internacionales en su campo. Ha sido miembro del comité organizador de 19 congresos internacionales y es miembro del comité editorial de 5 revistas internacionales.

El trabajo de Nuria ha sido reconocido internacionalmente con numerosos premios. Fue número uno de su promoción en la Escuela Superior de Telecomunicaciones de Madrid, y Premio Nacional de Telecomunicaciones (1994). Fue la primera española es recibir el MIT TR100 (hoy TR35) Young Innovator Award (2004) y el Rising Talent Award por el Women's Forum for the Economy and Society (2009). Ha recibido el Premio Ada Byron a la mujer europea digital del año (2016), la medalla al Mérito Empresarial y Social de la Generalitat Valenciana (2017), el premio EVAP a la Diversidad (2017), el Premio Nacional de Informática Angela Ruiz Robles (2016) y el Premio Ingeniera del Año (2018) entre otros.

Ha sido nombrada una de las 11 personas más influyentes de Inteligencia Artificial en el mundo por Pioneering Minds (2017), una de las mentes maravillosas españolas en tecnología por el periódico EL PAÍS (2017), “una destacada directora de tecnología” (EL PAÍS, 2012), uno de los “100 líderes para el futuro” (Capital, 2009) y uno de los “40 jóvenes que marcarán el próximo milenio” (EL PAÍS, 1999), entre otros.

Es inventora de 40 patentes. Su trabajo ha contribuido a la mejora de servicios, a la creación de servicios nuevos, a la definición de estrategias y a la creación de empresas nuevas. Nuria es ponente invitada regularmente en congresos internacionales tanto científicos como tecnológicos.

Además de su trabajo científico, Nuria dedica parte de su tiempo a la divulgación científico-tecnológica y a inspirar a jóvenes y adolescentes —y en especial a las chicas— a estudiar carreras técnicas. Ha impartido charlas a más de 8.000 adolescentes, ha contribuido con el capítulo titulado “Eruditos Digitales” para el libro “Los nativos digitales no existen” (Deusto, 2017), ha escrito artículos para EL PAÍS, The Guardian, TechCrunch entre otros y ha sido co-organizadora del primer evento de TEDxBarcelona dedicado a la educación emergente, entre otras actividades. Sus charlas de WIRED, TEDx y eventos similares han sido vistas miles de veces.

Twitter: @nuriaoliver  
nuria@alum.mit.edu

→ gtt  
gestión tributaria  
territorial