INTELIGENCIA ARTIFICIAL: PRESENTE Y FUTURO

Esteban García-Cuesta
Escuela de Arquitectura, Ingeniería y Diseño
UNIVERSIDAD EUROPEA DE MADRID

Hoy en día la tecnología está reflejada en muchos aspectos de nuestras vidas y es uno de los más importantes catalizadores de la transformación en la industria. De entre ellas la Inteligencia Artificial aparece como una de las áreas científicas que más impacto está teniendo en esta transformación y con relevantes casos de éxito. Esta transformación no sólo implica la generación de nuevos productos o servicios sino también un cambio de la sociedad del producto a la sociedad del conocimiento y de la interacción hombre-máquina. Este artículo expone los aspectos más relevantes que están propiciando este cambio junto con las tendencias y retos que surgen durante su aplicación.

PALABRAS CLAVES •
inteligencia artificial, ética, sociedad, computación, tecnología

CÓMO CITAR ESTE ARTÍCULO •
García-Cuesta, Esteban. 2020. "Inteligencia Artificial: presente y futuro" en :: UEM STEAM Essentials
Enlace web UEM :: http://projectbasedschool.universidadeuropea.es/escuela/escuela/steam_essentials

"Estudia el pasado si quieres pronosticar el futuro"

Confucio

INTRODUCCIÓN

La Inteligencia Artificial (IA) ha estado siempre entrelazada a la psicología cognitiva con el objetivo de replicar las capacidades humanas para resolver problemas y dotar a las máquinas con habilidades que les permitan la comprensión de su entorno. La comprensión normalmente se asienta en la caracterización del problema como una combinación de inferencias simbólicas y sub-simbólicas, la extracción de un conjunto de patrones, y el uso de métodos de aprendizaje automático. Aunque en los años 40 e inicios de los 50 hubo algunos trabajos incipientes sobre modelos de aprendizaje basados en redes neuronales artificiales (Mculloch & Pitts, 1943), se desarrollaron los primeros juegos de ajedrez (Shannon, 1950), y hubo trabajos de corte filosófico como el test de Turing (Turing, 1950) que redefinían el modo de abordar el término de inteligencia computacional, la concepción actual de la IA se le atribuye a las jornadas de trabajo que se realizaron en Dartmouth en 1956 (McCarthy *et al.*, 1955). En estas jornadas un grupo de profesores y científicos se reunieron con el objetivo de definir qué elementos debería tener una máquina para que pudiera replicar cualquier aspecto relacionado con el aprendizaje o cualquier característica de la inteligencia pudiendo así resolver problemas que hasta entonces estaban reservados solo para humanos. Esto incluía la necesidad del desarrollo de nuevos lenguajes máquina, el tratamiento y procesamiento de lenguaje natural, nuevos modelos de aprendizaje automático como las redes neuronales artificiales, teorización de la computación, y la abstracción máquina.

figura 01 » Biblioteca Baker, Universidad de Darmouth **(Gaving Huang CC BY 3.0)**

Durante los siguientes años (1952-1969) hubo un gran entusiasmo y grandes esperanzas por el éxito de estos sistemas inteligentes, tanto en la universidad abordando problemas sólo resolubles hasta entonces por humanos como por la industria que veía en la automatización de procesos y sistemas de decisión grandes oportunidades de aumentar su eficiencia. También se desarrollan nuevos lenguajes de programación de alto nivel como LISP (McCarthy *et al.*, 1960) que facilitaban la definición simbólica de hechos a través de listas, y la definición de reglas lógicas y de asociación. Esto permitía realizar inferencias lógicas y simular el razonamiento humano lo que hizo que la comunidad de IA lo adoptara rápidamente. Tras una primera década de entusiasmo, entre 1966-1973 se empezaron a entrever algunos problemas y limitaciones de los magnánimos objetivos definidos para la IA entrando así en lo que se denominó el primer invierno de la IA. Estas limitaciones incluían las puramente computacionales (para que el lector vea las características computacionales de la época, el IBM704 que fue uno de los primeros super-ordenadores de uso comercial y tenía 32KB de memoria o una doscientosava parte de un ordenador personal actual) y también algunos fracasos iniciales en el área de traducción automática.

Poco después hay un resurgimiento centrado especialmente en los sistemas inteligentes basados en conocimiento (1969-1979). Este resurgimiento se debe al éxito del uso de estos sistemas en el ámbito médico y de control de plantas industriales donde el conocimiento de los expertos se modela a través de esquemas simbólicos que recogen reglas de actuación y permitían representar el conocimiento haciendo posible la automatización de algunos procesos de decisión. El uso de ontologías y redes bayesianas para modelar estructuras de causa-efecto y el aprendizaje de modelos de decisión en base a esas estructuras se aplican con éxito a diversas aplicaciones mostrando así su enorme potencial. Entre estos sistemas cabe destacar el proyecto MYCIN (Buchanan, 1984) que permitía identificar bacterias causantes de infecciones y recomendar antibióticos con una dosis ajustada por paciente y peso. También surgieron nuevos lenguajes específicos en el área como Prolog en 1972 (Colmerauer, 1993) diseñado para permitir el uso de la lógica de primer orden y lógica formal a través de la expresión de relaciones con reglas y hechos.

Tras esta década de éxito en dominios aislados, es a partir de los 80 cuando se extiende el uso de la IA a otros ámbitos industriales en lo que se denominarían "Sistemas Expertos". Estos sistemas reproducen y simulan el conocimiento y algunas habilidades analíticas humanas como la inferencia para resolver problemas que son suficientemente complicados como para requerir un humano experto para su resolución (Feigenbau, 1980). Durante esta época comienzan también a asentarse nuevos aspectos de la IA que en la actualidad siguen vigentes como la minería de datos o la tecnología semántica, aunque no es hasta la década de los 90 cuando se desarrollan. A finales de este periodo da comienzo el segundo y último invierno de la IA propiciado por el colapso de las máquinas optimizadas para el uso del lenguaje LISP y el caro mantenimiento de los sistemas expertos desarrollados como XCON (McDermott, 1980). Esto provocó un colapso en un sector grande de la industria asociada a la IA quien veía en este lenguaje y en sus aplicaciones su futuro.

Es durante los años 90 cuando se comienza a usar de un modo relevante la IA para realizar minería de datos en áreas tan relevantes como la medicina (p.ej. diagnósticos

médicos) o la logística. A su vez el aumento significativo de las capacidades de cómputo que se duplicaban cada año (Ley de Moore; Moore 1965,1975) permitió establecer nuevos lazos con otras áreas como la estadística, la economía y la matemática haciendo que los investigadores se comprometieran a su vez con los métodos científicos estandarizados. En mayo de 1997 Deep Blue fue el primer sistema de IA en lograr batir al campeón del mundo Gary Kasparov en el juego de ajedrez. A nivel técnico se desarrollan métodos de aprendizaje automático muy relevantes como las máquinas de vector soporte (Boser, *et al.*, 1992) que han sido durante más de una década uno de los métodos con más éxito en problemas de clasificación y regresión, y en especial cuando se dispone de datos de alta dimensionalidad.

En paralelo durante las décadas de los 80 y 90 también se desarrolla el concepto actual de Internet y la Web; inicialmente como un conjunto de máquinas interconectadas y accesibles a través de hiperenlaces y posteriormente como la Word Wide Web (WWW) definiendo un conjunto de protocolos que permiten la comunicación de los ordenadores por Internet (Berners, 1990). Aunque la evolución de las comunicaciones no está ligada intrínsecamente a la IA, es cierto que desde la concepción de la WWW y con el aumento de las capacidades de comunicación han surgido muchas empresas y nuevos sectores puramente digitales que han ejercido una gran influencia en el área desde entonces. Ese es el caso de empresas como Google, Facebook, o Twitter por mencionar algunas y, como veremos a continuación, han logrado algunos hitos reseñables en el área de la IA actual.

Desde el comienzo del S.XXI se han continuado los desarrollos iniciados durante los 90 y revitalizado algunas áreas como las redes neuronales artificiales que están logrando tener una segunda juventud a través de la sinergia entre capacidades de cómputo y complejidad de modelo logrando explotar de manera singular las habilidades de generalización que poseen de manera intrínseca. Es destacable en este aspecto resaltar los grandes logros en visión por computador donde ya se utiliza para la conducción autónoma de coches. Pero no es hasta la última década donde podemos ver una nueva carrera de la IA por lograr superar nuevos grandes retos.

En el año 2011 en *"Jeopardy! Quiz show"* un sistema experto inteligente desarrollado por IBM y basado en su herramienta Watson ganó a dos de los más célebres concursantes por un amplio margen (Markoff, 2011). En los siguientes dos años Google y Facebook liberan sus motores de búsqueda Web (Knowledge Graph y Graph Search). Estos motores se basan en el análisis de relaciones y en análisis semántico de los miles de millones de páginas Web que están accesibles a través de la WWW para lograr búsquedas de información precisas, y su liberación como código abierto facilita su adopción por la industria y el rápido avance de muchas compañías del sector. Con el objetivo de poder replicar el comportamiento del cerebro tanto en Estados Unidos como en Europa se financian grandes proyectos para entender mejor como se establece la representación de los conceptos y su interrelación a nivel de sinapsis y localización cerebral, y surge el intento de replicar la funcionalidad del cerebro humano y de sus cerca de mil trillones de conexiones a través de una arquitectura hardware (SpiNNaker; Furber, 2014). Esta arquitectura emula el cerebro basándose en redes de neuronas artificiales e incorporando el concepto de tiempo además de neuronas y conexiones sinápticas como se había hecho hasta ahora. Aun así, esta arquitectura sólo es capaz de emular un uno por ciento de la escala del cerebro humano pareciéndose más al cerebro de un ratón.

Sin duda la última década ha sido una época de oro para la IA pero no se puede finalizar un repaso a su historia sin su último gran avance. En marzo de 2016 AlphaGo (Silver, *et al.* 2016), un sistema inteligente para jugar al Go, ganó 4 de las 5 partidas al campeón del mundo Lee Sedol siendo el primer sistema en batir a un jugador profesional de Go estableciendo así un nuevo hito. Además, este sistema inteligente es capaz de aprender desde cero a jugar sin necesidad de tener partidas con conocimiento previo como sucedía con Deep Blue. Sin duda, esta característica hace mucho más versátil a este tipo de sistemas donde se pueden generar datos de aprendizaje a partir de las propias reglas del juego o del contexto de aplicación. Modificaciones de este sistema han sido posteriormente utilizadas con éxito para ganar en el póker (Brown, 2019) o en juegos de estrategia (Berner, 2019).

Es importante resaltar que durante las más de cinco décadas de desarrollo de la IA cada una de estas áreas ha tomado su propia identidad y se han desarrollado trabajos relevantes en todas ellas muchas veces de manera independiente sin tener en cuenta los objetivos originales por las cuales surgieron. En gran medida es normal que esto suceda ya que al tratar en profundidad suficiente para resolver los enigmas de cada una de esas partes de la IA se pierde la visión global, "los árboles no dejan ver el bosque". Es por esto que actualmente se están abordando nuevos retos relacionados con los objetivos iniciales de la IA y también simultáneamente se están aplicando los conocimientos adquiridos en las últimas décadas especialmente a diferentes sectores de la sociedad.

ÁREAS DE APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL

Las grandes áreas de aplicación de la IA pueden clasificarse en: representación del conocimiento, aprendizaje automático, y sistemas inteligentes artificiales

Representación del conocimiento

Es el área de la IA dedicada a representar la información acerca del mundo de un modo que pueda ser utilizado por un sistema de computación para la resolución de tareas complejas como diagnósticos médicos, el diálogo hombre-máquina, o la predicción de eventos. **(ver tabla 01)**

Aprendizaje automático

Es el área de la IA que estudia los algoritmos que mejoran automáticamente a través de la experiencia (Mitchell, 1997). Estos algoritmos construyen modelos matemáticos basados en un conjunto de datos, denominados "datos de entrenamiento" para realizar predicciones o tomar decisiones sin ser programados de un modo explícito. Dentro de esta área se diferencian tres aproximaciones dependiendo de la naturaleza de la retroalimentación que tienen los modelos durante el aprendizaje.

SUB-ÁREA	DESCRIPCIÓN	OBJETIVOS
Solucionador general de problemas	Enunciado por Herbert A. Simon y J.C. Shaw en 1959 puede resolver problemas que son suficientemente formalizados (p.ej. Torres de Hanoi) pero inaplicable en problemas reales debido a la explosión combinatoria del espacio de soluciones. Su uso junto con técnicas de optimización de búsquedas (p.ej. A* o IDA) y heurísticas permiten abordar estos problemas más complejos con cierto éxito. Más recientemente el uso de búsquedas estocásticas junto con refuerzo han mostrado ser mucho más eficientes (Silver, D. *et al.* 2016).	Obtener una máquina capad de resolver cualquier problema que pueda ser expresado por formulación o clausular de Horn, o más recientemente por el aprendizaje por refuerzo.
Sistemas expertos	Requiere de formalismos, automatismos de razonamiento lógico, e incorporar conocimiento de dominio. Estos sistemas se dividen en dos partes: i) la base de conocimiento que representa los hechos conocidos sobre el mundo y ii) las reglas y motores de inferencia que representan las interacciones del mundo y se aplican sobre esa base de conocimiento. Por su especial importancia resaltar el uso de esta tecnología para crear lenguajes basados en cuadros de contexto (Subirats, *et al.* 2003) que permiten agrupar contextos lingüísticos con el fin de facilitar la interacción natural entre hombre-máquina.	Obtener sistemas especializados capaces de resolver de manera autónoma problemas concretos sobre dominios de aplicación.
Razonamiento con sentido común	La carencia por parte de una máquina del conocimiento base asociado a asunciones que se hacen y modos de comportamiento de las personas concluye en una gran limitación para la interacción hombre-máquina y la toma de decisiones. Para poder superar esta barrera se definen miles de conceptos y sus relaciones a través de los diferentes roles con los que pueden actuar en diferentes momentos. El proyecto más ambicioso en este sub-área ha sido Cyc (Lenat et al. 1986). que pretendía codificar, en un modo que fuera entendible por una máquina, los millones de piezas de conocimiento que componen el sentido común. En 2002 se liberó la primera versión OpenCyc. Actualmente gran parte de los nuevos trabajos son realizados en el ámbito de la Web Semántica construyendo ontologías de conceptos sobre la Web para enriquecer las búsquedas y facilitar el acceso al conocimiento.	Representar el sentido común de los humanos a través de cuadros contextuales de conocimiento sobre física, la temporalidad, la causalidad, intenciones y otros (Lenat *et al.*. 1986).

tabla 01 » Sub-áreas de representación del conocimiento

1 » Aprendizaje supervisado: el computador aprende a partir de ejemplos y de las salidas deseadas para cada uno de ellos siendo el objetivo aprender el modelo o regla que asocia las entradas con las salidas. Algunas fde las técnicas más exitosas de este tipo son las máquinas de vector soporte, *AdaBoost*, *RandomForest*, o la redes neuronales artificiales.

2 » Aprendizaje no supervisado: ante la ausencia de las salidas deseadas tiene como objetivo descubrir la propia estructura de los ejemplos. Puede servir para descubrir patrones ocultos en los datos que faciliten su interpretación o ser un paso previo para el aprendizaje de las características más relevantes en un proceso de generación de conocimiento. Algunas de las técnicas más utilizadas son k-means, los mapas auto-organizados, análisis de componentes principales, o las redes generativas antagónicas.

3 » Aprendizaje por refuerzo: el algoritmo interactúa directamente con un entorno dinámico en el que debe alcanzar un objetivo (p.ej. ganar una partida de ajedrez). Durante la interacción el algoritmo navega por el espacio de posibles soluciones obteniendo recompensas para cada una de las acciones tomadas reforzando aquellas que proporcionan un mayor beneficio frente al resto. Como resultado se obtienen unas reglas de actuación que maximizan la probabilidad de conseguir el resultado deseado. Algunas de las técnicas más utilizadas son el método Monte Carlos, Q-Learning, o DQN.

A menudo los términos de minería de datos y aprendizaje automático son utilizados de manera equivalente, pero aunque comparten ciertos elementos y técnicas su principal diferencia está en que el primero trata de descubrir nuevo conocimiento (tiene un carácter más exploratorio) mientras que el segundo trata de mejorar el rendimiento sobre conocimiento conocido.

Sistemas Inteligentes Artificiales e Inteligencia Artificial General

La Inteligencia Artificial General busca un modo de aprender y modelar el mundo para resolver cualquier tipo de problema. Debido a la complejidad inherente a la interconexión de dominios y conceptos se busca, adaptar esta inteligencia artificial general a cada uno de ellos con la intuición que puedan posteriormente unirse formando un único sistema. Estos Sistemas Inteligentes Artificiales específicos incluyen muchas de las técnicas y metodologías de otras áreas y se especializan en dominios concretos para lograr aprender y solucionar problemas complejos, pero en un contexto bien definido. A pesar del gran interés y de los distintos intentos realizados, todavía no se ha desarrollado un sistema que se considere generalista. Como ejemplo de estos intentos, existen sistemas avanzados capaces de aprender automáticamente interpretaciones semánticas de la información de la Web como "NELL - Never Ending Language Learning" (Carlson, et al. 2010) o capaces de almacenar conocimiento con ontologías semánticas como IBM's Watson con el objetivo de proporcionar sistemas inteligentes artificiales de pregunta-respuesta con una base de conocimiento permanentemente actualizada y veraz.

Desde otra perspectiva, no tan centrada en la adquisición y representación del conocimiento sino en la optimización de la búsqueda en espacios de dimensiones gigantescas, los algoritmos desarrollados por DeepMind como AlphaGo han demostrado también su capacidad de lograr aprender en diferentes situaciones las mejores estrategias para poder batir a un oponente sin que ello implique un cambio de la algoritmia que tienen implementada (Silver, 2018). Esto también refleja en gran medida la capacidad de aprender de manera genérica y muestra rasgos de inteligencia artifical general a través de las acciones tomadas tras dicho aprendizaje.

SECTORES DE APLICACIÓN DE LA IA

La IA moderna es ubicua y un listado exhaustivo de sus aplicaciones sería inviable por su numerosidad. Por esto se han agrupado en cuatro sectores relevantes para realizar una descripción de algunos de los logros en cada uno de ellos. **(ver tabla 02)**

RETOS Y TENDENCIAS

Después de más de 60 años, se puede afirmar que los grandes retos de la IA siguen siendo los que hicieron el surgimiento de esta disciplina como área de investigación y que su máximo exponente es la Inteligencia Artificial General la cual se ha visto desplazada durante muchos años por la dificultad que entrañaba y la poca experiencia de los inicios. Para lograr ese fin se han desgajado durante estos años varios objetivos parciales que han creado varias tendencias actuales.

Por un lado, se dispone de modelos de aprendizaje automático muy sofisticados y especializados en diferentes aplicaciones de dominio donde obtienen unos resultados sorprendentes debido a la eficiencia que se ha logrado en el desarrollo de estas técnicas. Esto se ve reflejado en su máximo exponente en los algoritmos que integran dichos modelos junto con métodos de búsqueda en espacio de estados donde se complementan mutuamente. Éste segundo permite ante la carencia de información de dominio o experiencia previa explorar cualquier posible solución evitando así el descarte por desconocimiento de posibles soluciones

CIENCIAS DE LA SALUD	El mayor uso de la IA en salud ha sido como método de clasificación automática para el análisis de evidencias como imágenes de resonancia magnética o señales de electrocardiogramas (ECG) en pacientes con riesgo de enfermedad. El éxito de su uso como sistemas de ayuda a la decisión médica ha provocado que se extienda también a otro tipo de aplicaciones en este ámbito que pueden redundar en el descubrimiento de nuevos tratamientos no conocidos (Fleming, 2018), explicación de algunos procesos asociados a una enfermedad (Yu, et al. 2017), o detección temprana de enfermedades a partir de historiales clínicos (Mesón, R. & García-Cuesta, E. 2017). También en el ámbito de la biomedicina ha sido aplicado con éxito para identificar la asociación de genes con enfermedades y con respuestas farmacológicas (Lindeboom, 2019).
ROBÓTICA	La robótica y la IA están entrelazadas por la relación de la parte cognitiva y la parte física las cual permite interaccionar a los algoritmos con el mundo real mecanizado. La interacción es dual, por un lado los sensores permiten recolectar información del mundo mientras que a través de las acciones robóticas se pueden materializar las decisiones tomadas por la IA. Algunos de los usos más relevantes son en el campo de la conducción automática con el desarrollo de los coches autónomos (todavía en fase de test) capaces de identificar en tiempo real con algoritmos de visión por computador la escena que se presenta. También tiene aplicabilidad en otros contextos como la asistencia en casa a personas con Alzheimer proveyéndoles información sobre sus rutinas y posibles despistes, o en el aprendizaje inclusivo facilitando la interacción de niños con síndrome del espectro autista (INCLUDED, 2019).
SOCIEDAD Y ECONOMÍA DIGITAL	La IA tiene varias formas de interaccionar con la sociedad. Por un lado, facilita el acceso a información haciéndola más democrática llegando a la personalización según las necesidades, y por otro dinamiza y hace más eficiente el proceso económico. En economía se realizan predicciones de la evolución de los mercados pudiendo introducir eventos e información externa en tiempo real mejorando así su rendimiento (en el 2001 una máquina logra batir a un grupo de personas en una competición simulada), o se puede detectar el fraude por el uso indebido de tarjetas fomentando así adopción de una economía digital entre otros. El derecho de expresión y el acceso a la información son aspectos de gran importancia para lograr una sociedad entre iguales. En este contexto la IA permite a través del análisis e interpretación semántica de los contenidos Web facilitar éste acceso a cualquier persona en cualquier momento convirtiéndose así en una herramienta promotora de la equidad social.
INDUSTRIA 4.0	La IA está siendo uno de los más importantes catalizadores en la trasformación digital y sin duda es uno de los que más repercusiones tendrá en el futuro por su aportación al incremento de valor añadido en la cadena de producto o servicio. Esta cuarta revolución industrial se caracteriza por los sistemas inteligentes y autónomos los cuales son capaces de analizar grandes volúmenes de datos y tomar decisiones en tiempo real en base a modelos aprendidos y los datos. En este contexto la IA se está utilizando para hacer más eficiente algunos procesos industriales ya existentes. Es el caso del mantenimiento de maquinaria y equipos a través de la predicción de fallos lo cual permite la reparación preventiva (Lee, 2019) o la mejora del proceso de manufactura en lo que se ha denominado la "manufacturación inteligente" siendo capaces de regular y adaptar todo el proceso a las necesidades del cliente (Lu, 2016). Estos son algunos ejemplos, pero el cambio es ubicuo a la industria y afectará tanto a la parte organizacional como a la operacional y de producto.

tabla 02 » Descripción de sectores más relevantes de aplicación de la IA

óptimas, y el modelo de aprendizaje automático permite predecir como de buena será una acción tomada durante los primeros estadios del proceso de búsqueda sin tener que explorar completamente todo el espacio lo cual sería extremadamente costoso (pudiendo llevar incluso años la realización de dicha exploración completa) y carente de utilidad real.

Por otro lado, tenemos la explosión de información que nos ha proporcionado la Web y el surgimiento de la Web semántica como un intento de interpretar y generar conocimiento a partir de dicha información. Es en este contexto donde toda la tecnología de representación del conocimiento y del procesamiento de lenguaje natural se unifican para generar la Web del conocimiento la cual sea fuente para los métodos de aprendizaje IA y ayude a mejorar procesos de toma de decisiones, generar comunidades y facilitar el acceso al conocimiento, y permitir a las máquinas la interacción con este conocimiento de manera automática pudiendo adaptar de manera constante su comportamiento a los cambios que suceden. En este contexto proyectos pioneros como Never Ending Languaje Learning marcan algunos de los resultados esperados sobre los cuales pueden diseñarse nuevas aplicaciones.

Actualmente dada la gran eficiencia lograda en cada una de estas áreas a las que (Domingos, 2015) denomina "tribus", se intuye que cada una tienen un gran potencial de contribución para lograr una única Inteligencia Artificial General definiendo así el denominado "algoritmo maestro".

Estas tendencias no quedarían completas si no se mencionara uno de los aspectos adjuntos al propio desarrollo de la IA como es la ética. Como consecuencia de la integración cada vez más habitual de este tipo de sistemas inteligentes artificiales en nuestra vida cotidiana surge una preocupación cada vez mayor por saber si esta interacción entre personas y algoritmos (o máquinas) la cual tiene una gran diversidad de formas (p.ej. a través de las búsquedas en internet, la recomendación de noticias en aplicaciones, o de las rutas más óptimas en el transporte) está siendo adecuada según los criterios morales, éticos y legales de nuestra sociedad. Esto da lugar a diversas opiniones y corrientes tanto a favor como en contra teniendo las primeras su máxima expresión en el transhumanismo y las segundas un grupo de resistencia que, aunque actualmente no tiene una forma determinada, es inherente a toda sociedad ante cambios bruscos o carentes de la adecuada pedagogía social. Como un primer paso hacia ese proceso de pedagogía se ha establecido una guía a nivel Europeo donde se recogen los diferentes aspectos a tener en cuenta para establecer una IA confiable (EU, 2019). Este primer paso refleja los aspectos a tratar, pero no la descripción de una solución ni la automatización de esta IA confiable. Para ello también es necesario el desarrollo de entornos software donde se puedan adherir dichos sistemas inteligentes artificiales y ser verificados éticamente promoviendo el concepto de una "IA ética por diseño" (García-Cuesta, 2020b).

CONCLUSIÓN

La inteligencia artificial después de diferentes periodos de auge y decaimiento por los que ha pasado durante lo más de 60 años de existencia está en un momento álgido donde confluyen varios elementos que pueden propiciar su próximo gran avance. En el ámbito industrial se ha demostrado que tiene una gran importancia en la optimización de procesos, la generación de valor añadido, y en la creación de nuevas aplicaciones e incluso de nuevos sectores económicos completos. Así mismo en el ámbito de la investigación se ha culminado con gran éxito una etapa de más de 20 años de división de los grandes retos en otros más pequeños y asequibles. En este periodo, se ha logrado disponer de sistemas inteligentes artificiales específicos con aplicaciones que abarcan desde la detección automática de enfermedades hasta el aprendizaje automático sin conocimiento previo logrando superar a jugadores profesionales en juegos que hasta ahora se consideraba imposible que una máquina ganara. Estos dos sucesos propician que sea un momento bueno para que las dos grandes áreas de la IA, la representación del conocimiento y el aprendizaje automático, confluyan para abordar la Inteligencia Artificial General. Sin duda es el momento, pero no estará exento de problemas y dudas por lo que serán necesarias grandes mentes que enseñen a aprender.

REFERENCIAS BIBLIOGRÁFICAS

» Berners-Lee, T.(1990) *"Information Management: A Proposal"*. CERN.

» Berner, C. et al. OpenAI. (2019) *"Dota 2 with Large Scale Deep Reinforcement Learning"*. arXiv:1912.06680 [cs.LG] arXiv:1912.06680

» Noam Brown, Tuomas Sandholm. (2019) "*Superhuman AI for multiplayer poker"*. Science, 2019; eaay2400 DOI: 10.1126/science.aay2400

» Buchanan, B. G. & Shortliffe, E.H. (1984) *"Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project"*. Addison Wesley, Reading, MA.

» A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E.R. Hruschka Jr. and T.M. Mitchell. *"Toward an Architecture for Never-Ending Language Learning"*. In Proceedings of the Conference on Artificial Intelligence (AAAI), 2010

» Domingos, P. (2015) *"The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World"*, Publisher: Basis Books, ISBN 978-0465065707

» EU (2019) High-Level Expert Group on Artificial Intelligence *"Ethics Guidelines for Trustworthy AI"*. European Commision B-1049 Brussels, 8 April 2019.

» E A Feigenbaum. (1980) *"EXPERT SYSTEMS IN THE 1980s"*. Annual Report to the National Institutes of Health

» Boser, Bernhard E.; Guyon, Isabelle M.; Vapnik, Vladimir N. (1992). "*A training algorithm for optimal margin classifiers"*. Proceedings of the fifth annual workshop on Computational learning theory – COLT '92. p. 144.

» Fleming N. *"How artificial intelligence is changing drug discovery"*. Nature. 2018;557(7707):S55-S57. doi:10.1038/d41586-018-05267-x

» Furber, S. B.; Galluppi, F.; Temple, S.; Plana, L. A. (2014). *"The SpiNNaker Project"*. Proceedings of the IEEE. 102 (5): 652–665.

» García-Cuesta, E., López-López, José M., Gómez-Vergel, D., Huertas-Tato, J. (2020a) *"An Adaptive Cognitive Model to Integrate Machine Learning and Visual Streaming Data"*. 15th International Conference on Soft Computing Models in Industrial and Environmental Applications SOCO'2020, Burgos

» García-Cuesta, E. (2020b) *"Artificial Intelligent Ethics in the Digital Era: an Engineering Ethical Framework Proposal"*. arXiv:2002.07734

» INCLUDED (2019) *"Digital Storytelling for Inclusion"* EU Project funded by Erasmus + EU programme (2019-2022), **www.includedeurope.eu** (last visited 12/07/2020)

» Lindeboom, R.G.H., Vermeulen, M., Lehner, B. et al. (2019) *"The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy"*. Nat Genet 51, 1645–1651. **https://doi.org/10.1038/s41588-019-0517-5**

» Lee, W.J. et al. (2019) *"Predictive Maintenance of Machine Tool Systems Using Artificial Intelligence Techniques Applied to Machine Condition Data"*, Procedia CIRP Volume 80, 2019, Pages 506-511

» Lenat, Doug; Prakash, Mayank; Shepherd, Mary (1986). *"CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquistion [sic] Bottlenecks"*. AI Magazine. 6 (4): 65–85. ISSN 0738-4602.

» Lu, Y., Morris, K.C., and Frechette, S. (2016) *"Current Standards Landscape for Smart Manufacturing Systems,"* National Institute of Standards and Technology.

» Markoff, John (16 February 2011). *"Computer Wins on 'Jeopardy!': Trivial, It's Not"*. The New York Times. Retrieved 28 June 2020

» McCarthy, J.; Brayton, R.; Edwards, D.; Fox, P.; Hodes, L.; Luckham, D.; Maling, K.; Park, D.; Russell, S. (1960). *"LISP I Programmers Manual"*. (PDF). Boston, Massachusetts: Artificial Intelligence Group, M.I.T. Computation Cente and Research Laboratory.

» McCulloch, W. S., & Pitts, W. (1943). *"A logical calculus of the ideas immanent in nervous activity"*. The bulletin of mathematical biophysics, 5(4), 115-133.

» McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955), *"A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence"*.

» McDermott, John (1980). *"R1: An Expert in the Computer Systems Domain"*. Proceedings of the First AAAI Conference on Artificial Intelligence. AAAI'80. Stanford, California: AAAI Press: 269–271.

» Mesón, R. & García-Cuesta E. "*Sistema de predicción de enfermedades basado en resultados históricos de pacientes: un estudio experimental"* Tesis Master Universitario en Análisis de grandes cantidades de datos - MBI. Biblioteca CRAI, Universidad Europea de Madrid.

» Mitchell, T. (1997) *"Machine Learning"*. McGraw Hill, 1997.

» Moore, Gordon E. (1965). *"Cramming more components onto integrated circuits"*. Electronics Magazine. p. 4. .

» Moore, Gordon (1975). *"IEEE Technical Digest 1975"*. Intel Corp.

» Muñoz, R. & García-Cuesta, E. *"Big data architecture for prediction and decision taking in clinical environments"*. The 5th Big Data Spain Conference 17-18 Nov. Madrid, Spain

» Silver, D., Huang, A., Maddison, C. et al. (2016) *"Mastering the game of Go with deep neural networks and tree search"*. Nature 529, 484–489. **https://doi.org/10.1038/nature16961**

» Shannon, C. (1950). *"Programming a Computer for Playing Chess"*. Philosophical Magazine. 41 (314).

» Silver, D. et al. *"A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play"*. Science 07 Dec 2018: Vol. 362, Issue 6419, pp. 1140-1144 DOI: 10.1126/science.aar6404

» Subirats-Rggeberg, Carlos and Miriam R. L. Petruck (2003): *"Surprise: Spanish FrameNet!"* Presentation at Workshop on Frame Semantics, International Congress of Linguists. July 29, 2003, Prague, Czech Republic

» Turing, A. (1950). *"Computing Machinery and Intelligence, Mind"*, Volume LIX, Issue 236, October 1950, Pages 433–460, **https://doi.org/10.1093/mind/LIX.236.433**

» Yu, P., Wilhelm, K., Dubrac, A. et al. "FGF-dependent metabolic control of vascular development. Nature 545, 224–228 (2017). **https://doi.org/10.1038/nature22322**

BIOGRAFÍA

Esteban García-Cuesta es profesor titular en Ciencias de la Computación e Inteligencia Artificial del departamento de Ciencia, Computación y Tecnología de la Universidad Europea de Madrid y director del grupo de investigación en "Aprendizaje Automático y Ciencia de datos". Su actividad docente e investigadora ha estado vinculada desde su inicio al área de inteligencia artificial y las ciencias de la computación. Realizó estancias de formación investigadora en las universidades norteamericanas de Carnegie Mellon University (PA, US) en los años 2007, 2009 y 2010, y Rowan University en el 2008 donde realizó parte de sus estudios de investigación para la obtención de su doctorado (sobresaliente "cum laude") por la Universidad Carlos III de Madrid. A nivel formativo también posee diversas certificaciones empresariales relacionadas con el ámbito tecnológico y científico; y cuenta con cursos de gestión académica e investigación en LSE gracias a becas concedidas por la Fundación Ramón Areces. A nivel de industria obtuvo una beca competitiva Torres Quevedo y ha participado como jefe de proyectos y gerente de investigación en procesos de transferencia tecnológica. En el ámbito investigador está especializado en Computación e Inteligencia Artificial, cuenta con dos sexenios de investigación reconocidos por la ANECA, ha publicado más de 35 artículos científicos, es revisor nacional de proyectos de investigación, es revisor de revistas científicas de prestigio como MDPI Sensors o Applied Science, revisor de congresos internacionales, y miembro co-fundador de la asociación AI-Network para la promoción del uso de la Inteligencia Artificial en la industria y el bienestar social. Sus trabajos se han publicado en revistas y congresos científicos como International Journal of Neural Systems, Expert Systems With Applications, o International Semantic Web Conference, y participa activamente en la comunidad científica y tecnológica a través de seminarios, mesas redondas, publicaciones en periódicos, proyectos de investigación, y actividades de difusión tecnológica y emprendimiento.