

# DESARROLLO Y RETOS DE LA INTELIGENCIA ARTIFICIAL



Yuniesky Coca Bergolla  
Miguel Llivina Lavigne





# DESARROLLO Y RETOS DE LA INTELIGENCIA ARTIFICIAL

**Yuniesky Coca Bergolla**  
**Miguel Llivina Lavigne**



EDITORIAL EDUCACIÓN CUBANA

La Habana, noviembre de 2021

Esta obra está bajo una Licencia Creative Commons  
Atribución-NoComercial-CompartirIgual 4.0 Internacional.



© Yuniesky Coca Bergolla, 2021

© UNESCO, 2021

Diseño y Edición:

DI. Ismael Adán Quesada Chow

**ISBN: 978-959-18-1340-4**

Sello Editor EDUCACIÓN CUBANA

Dirección de Ciencia y Técnica - MINED

Calle 17 y O. Vedado. La Habana, Cuba.



# ÍNDICE

INTRODUCCIÓN .....	5
INICIOS Y DESARROLLO DE LA INTELIGENCIA ARTIFICIAL Y SU ENSEÑANZA .....	9
■ La enseñanza de la IA en Cuba .....	20
■ Futuro de la inteligencia artificial en el ámbito educativo .....	24
LA NECESARIA REFLEXIÓN ÉTICA .....	29
■ El debate sobre ética .....	33
■ En busca de consensos .....	35
CONSIDERACIONES FINALES .....	45
GLOSARIO DE TÉRMINOS .....	47
REFERENCIAS .....	51





# INTRODUCCIÓN

---

Este libro es una publicación científica, didáctica, metodológica y educativa, destinada a maestros, profesores, directores, metodólogos, inspectores, supervisores, técnicos, trabajadores de los distintos niveles educativos y otros profesionales de los organismos, instituciones y organizaciones gubernamentales y no gubernamentales en la República de Cuba y República Dominicana. Es el primero de cuatro libros que tienen como objetivo principal propiciar la incorporación de conceptos e ideas generales sobre Inteligencia Artificial en las escuelas, familias y comunidades. Este primer libro va dirigido a los fundamentos, historia y las cuestiones éticas que deben acompañar al desarrollo de esta ciencia.

El primer capítulo expone algunos de los fundamentos de la IA, específicamente las ciencias que han contribuido a su nacimiento y desarrollo. Además, se presentan algunos hechos relevantes que han marcado su desarrollo y su impacto en la sociedad. Se hace un apartado para su enseñanza en Cuba y se brindan algunos elementos de su futuro y retos en el ámbito educativo. El segundo capítulo presenta algunos de los riesgos que ya hoy amenazan a la humanidad por violar principios éticos en el desarrollo de sistemas con Inteligencia Artificial. Se presentan criterios y acciones específicas que se

proponen para lograr un desarrollo ético y sostenible de la Inteligencia Artificial. Finalmente, se pone a disposición de los lectores un Glosario de Términos y la Bibliografía.

Como antecedentes de esta publicación se elaboró en la Universidad de las Ciencias Informáticas (UCI) un documento base, el cual fue distribuido a especialistas, profesores, investigadores y otros profesionales relacionados con la Inteligencia Artificial. En él se brindan las ideas generales de posibles temáticas a abordar en edades tempranas, las cuales pudieran ser formalizadas en la educación general o tratadas como temas complementarios en diversos espacios. De igual forma, como parte del IV Taller Internacional de Enseñanza de las Ciencias Informáticas en el marco de la IV Conferencia Científica Internacional de la Universidad de las Ciencias Informáticas UCIENCIA 2021, se efectuó un panel de expertos donde participaron como panelistas, reconocidos profesores e investigadores de IA en Cuba. Todos los criterios fueron tomados en cuenta para enriquecer las ideas que aquí se exponen.

Aprovechamos la ocasión para expresar nuestro agradecimiento a todos los que aportaron opiniones, valoraciones y sugerencias para mejorar esta obra. Sobre todo al grupo de investigación de Inteligencia Artificial de la Universidad de las Ciencias Informáticas, al grupo de debate sobre ética y robótica de la Oficina Regional de la UNESCO en La Habana, integrado por desarrolladores y otros especialistas de los joven club, universidades y la propia oficina. De igual forma al colectivo de la disciplina Inteligencia Computacional de la carrera Ingeniería en Ciencias Informáticas de la UCI.

Esta obra pretende servir de base para contribuciones teóricas y prácticas mediante el desarrollo de actividades educativas curriculares y no curriculares para las instituciones educativas, escuelas, universidades, familias, comunidades y la sociedad en general, así

como para las instituciones y organizaciones gubernamentales y no gubernamentales; y de todos los sectores e instituciones de la sociedad. No se trata de enseñar a desarrollar aplicaciones de Inteligencia Artificial, sino, de contribuir al desarrollo del pensamiento computacional y nociones básicas necesarias para incentivar el estudio y profundización en IA con un posicionamiento ético, crítico y creativo, desde edades tempranas. Finalmente los exhortamos a leer, estudiar y profundizar en los aspectos que se abordan, para contribuir entre todos al desarrollo de la Inteligencia Artificial como componente esencial para la informatización de la sociedad.



---

# INICIOS Y DESARROLLO DE LA INTELIGENCIA ARTIFICIAL Y SU ENSEÑANZA

---

La Inteligencia Artificial (IA) es un área multidisciplinaria. Desde sus fundamentos hasta sus aplicaciones más recientes, abarca una amplia gama de sectores de la economía, la ciencia y la sociedad en su conjunto. Son varias las disciplinas que han contribuido con ideas, puntos de vista y técnicas, al desarrollo de la IA, al tiempo que la IA ha contribuido al desarrollo de ellas y otras muchas.

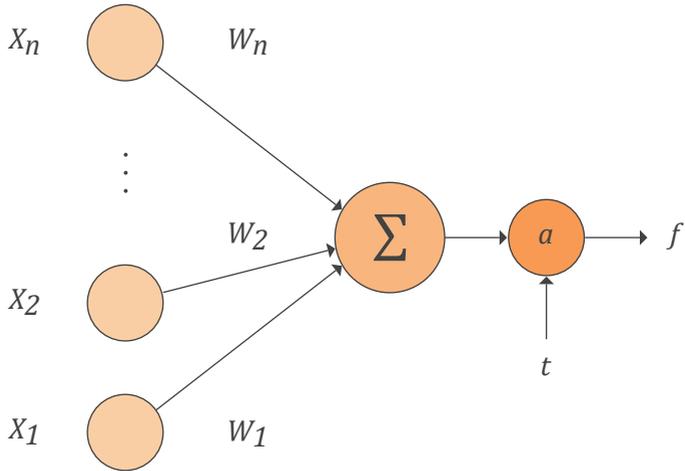
Para hablar de las bases de la IA debemos remontarnos al siglo III a. n. e. y analizar aportes teóricos y prácticos de varias ciencias (Russell & Norvig, 2004). Los filósofos concibieron la idea de que la mente es como una máquina que funciona a partir del conocimiento codificado en un lenguaje interno y el pensamiento servía para seleccionar las acciones a llevar a cabo. Las matemáticas proporcionaron herramientas para manipular las afirmaciones de certezas lógicas, las inciertas de tipo probabilista y prepararon el terreno para un entendimiento de lo que es el cálculo y el razonamiento con algoritmos. Los economistas formalizaron el problema de la toma de decisiones para maximizar los resultados esperados. Los psicólogos adoptaron la idea de que los humanos y los animales podían considerarse máquinas de procesamiento de información. Los lingüistas se encaminaron a demostrar que el uso del lenguaje se ajusta a ese modelo.

Los informáticos, por su parte, proporcionaron herramientas que hicieron posible la aplicación de la IA. Los métodos de IA tienden a ser extensos, al abordar problemas complejos no podrían funcionar sin los grandes avances en velocidad y memoria aportados por la industria informática. La teoría de control se centra en el diseño de dispositivos que actúan de forma óptima con base en la retroalimentación que reciben del entorno en el que están inmersos, una tarea que cada vez más asume la IA. Las herramientas matemáticas de la teoría de control eran bastante diferentes a las técnicas que utilizaba la IA, pero en la actualidad ambos campos se acercan cada día más (Russell & Norvig, 2004).

Los fundamentos de la IA datan de los silogismos de Aristóteles, presentados en el siglo III antes de nuestra era. Sin embargo, se considera el primer trabajo de Inteligencia Artificial la propuesta de red neuronal artificial diseñada por McCulloch y Pitts en 1943 (Figura 1). Ellos partieron de tres fuentes: conocimientos sobre la fisiología básica y funcionamiento de las neuronas en el cerebro, el análisis formal de la lógica proposicional y la teoría de la computación. Ellos sugirieron que redes adecuadamente definidas podrían aprender. Donald Hebb en 1949 propuso y demostró una sencilla regla de actualización para modificar las intensidades de las conexiones entre neuronas. Su regla, ahora llamada de aprendizaje Hebbiano, sigue vigente en la actualidad. Marvin Minsky y Dean Edmonds construyeron el primer computador a partir de una red neuronal en 1951.

Hay otro número de trabajos iniciales que se pueden considerar dentro del campo de la IA, pero se reconoce como fundador a Alan Turing quien articuló una visión de la IA en su artículo *Computing Machinery and Intelligence* (Turing, 1950). En él introduce la llamada prueba de Turing (Figura 2) para determinar si una máquina era

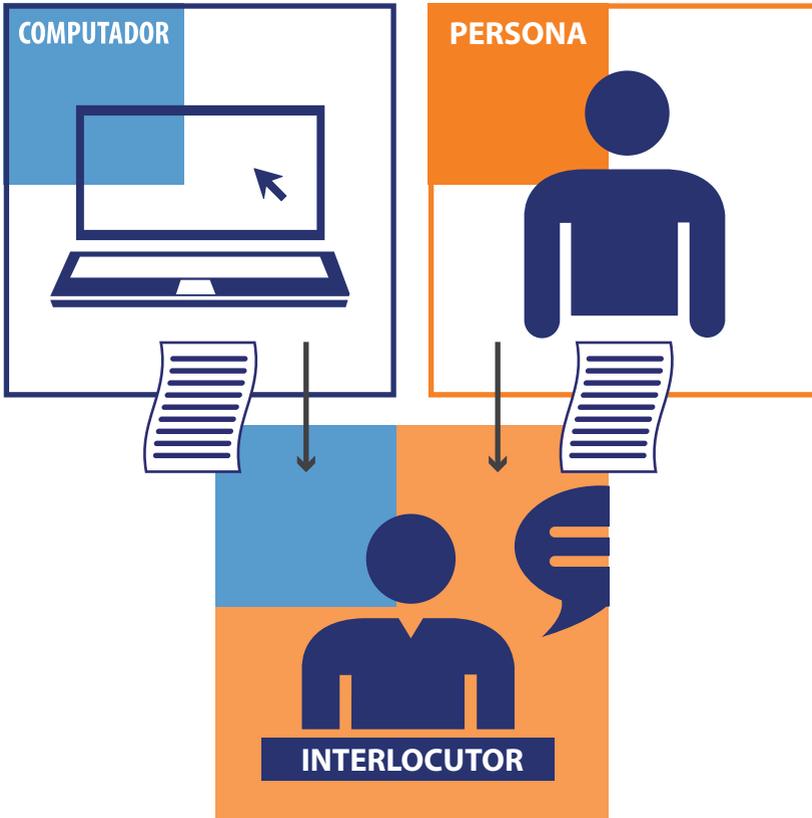
realmente inteligente y los elementos básicos del aprendizaje automático, los algoritmos genéticos y el aprendizaje por refuerzo.



$$f = 1, \text{ if } \sum_{i=1}^n x_i w_i > t$$

**Figura 1.** Estructura básica de una neurona artificial del modelo de McCulloch y Pitts.

En el verano de 1956 se organizó un taller con investigadores para debatir sobre sus experiencias y fomentar las investigaciones en el área de las redes neuronales, teoría de autómatas y de inteligencias en general. El principal resultado del taller fue la aceptación del término Inteligencia Artificial, propuesto por McCarthy. Aunque también se analizaron las características que la convierten en campo independiente, con problemas específicos que atender dentro de la rama de la informática.



**Figura 2.** La prueba de Turing propone cómo determinar si un sistema es inteligente.

En este contexto la *Association for Computing Machinery (ACM)*, reconocida como la primera sociedad científica y educativa para educar acerca de la Computación (ACM, 2020), incluye en su primera propuesta para currículos de Ciencias de la Computación (Atchison et al., 1968) un curso electivo de IA. Las temáticas que aborda se

reconocen como avanzadas para ese entonces e incluyen elementos teóricos básicos y de programación heurística.

Esos primeros años de la IA estuvieron llenos de éxitos. Uno de los principales resultados en esa etapa fue la creación del LISP, lenguaje funcional que se convertiría en el lenguaje de programación dominante en la IA y el segundo más antiguo que llega a nuestros días. Otro resultado importante fue en 1962 cuando aparecen las redes neuronales Adaline y Perceptrón. Estas fueron las primeras redes que podían entrenarse para resolver problemas más específicos.

Fue tanto el entusiasmo en la comunidad científica en esos primeros años, que se hicieron previsiones que nunca llegaron a cumplirse, por lo que desde finales de la década del 60 y toda la década del 70 fue de crisis en la IA. El financiamiento de los gobiernos para las investigaciones disminuyó significativamente y muchas investigaciones se dirigieron a demostrar la incapacidad de los algoritmos y métodos, empleados hasta ese momento, para resolver problemas verdaderamente complejos.

### Algunas de las principales causas que provocaron este estancamiento fueron (Russell & Norvig, 2004):

- **Los primeros programas contaban con poco o ningún conocimiento de la materia objeto de estudio.** Un ejemplo fue la traducción de texto. Se consideró que todo se reduciría a sencillas transformaciones sintácticas apoyadas en las gramáticas de los idiomas y al emplazamiento de palabras mediante un diccionario electrónico, lo que bastaría para obtener el significado exacto de las oraciones. La realidad es que, para traducir, es necesario contar con un conocimiento general sobre el tema.

- **No se utilizaban métodos para dar solución a los problemas intratables que intentaba resolver la IA.** La mayoría de los primeros programas de IA resolvían problemas experimentando con diversos pasos hasta que se llegara a encontrar una solución. Sin embargo, se comenzaron a encontrar problemas con explosión combinatoria donde la solución que proponían quedaba en manos del incremento de la velocidad del hardware. La utilización de heurísticas no era aún, una de las principales herramientas de la IA.
- **Limitaciones inherentes a las estructuras básicas que se utilizaban en la generación de la conducta inteligente.** Se demostró que si bien era posible lograr que los perceptrones (una red neuronal simple) aprendieran cualquier cosa que pudiesen representar, su capacidad de representación era muy limitada. En particular, un perceptrón con dos entradas no se podía entrenar para que aprendiese a reconocer cuándo sus dos entradas eran diferentes.

Por esa fecha aparecen los primeros sistemas basados en el conocimiento, como DENDRAL en 1969, el primero de ellos que tuvo éxito. Posteriormente aparece MYCIN para el diagnóstico médico, con tratamiento de incertidumbre mediante valores de certeza. El crecimiento generalizado de aplicaciones para solucionar problemas del mundo real provocó el respectivo aumento en la demanda de esquemas de representación del conocimiento que funcionaran. Se desarrolló una considerable cantidad de lenguajes de representación y razonamiento diferentes. Algunos basados en la lógica como el lenguaje Prolog que gozó de mucha aceptación años después, principalmente en Europa y Japón. Aún se mantiene como uno de los lenguajes representativos de la IA.

A raíz de la situación de crisis, en la próxima propuesta de la ACM (Austing, Barnes, Bonnette, Engel, & Stokes, 1979), la IA sigue siendo un curso electivo con elementos básicos para la profundización de los estudiantes. Se incorporan como elementos de mayor importancia, la representación del conocimiento, se abordan las estrategias de búsqueda como parte de esta representación y como parte de estructuras de control. Se incorporan las temáticas de comunicación, percepción y ejemplos de aplicaciones específicas.

En la década del 80, la IA se convierte en una industria (Russell & Norvig, 2004) gracias al auge de los sistemas expertos. En 1981 los japoneses anunciaron su proyecto de Quinta Generación, un plan de diez años para construir computadores inteligentes en los que pudiese ejecutarse Prolog.

Los modelos llamados conexionistas surgen de aplicar nuevos algoritmos de aprendizaje en las redes neuronales. Estos algoritmos de retroalimentación, utilizados en las redes multicapa, permitieron el gran resurgimiento de la investigación en redes neuronales de finales de los años 80, pero en realidad, se descubrieron por primera vez en 1969 por Bryson y Ho.

En el marco de este nuevo auge de la IA, en la propuesta de currículo de la ACM en 1991 (ACM-IEEE-CS, 1991) se incluye por primera vez la IA como parte del currículo básico y se incluyen elementos de historia, aplicaciones de la IA, problemas de estados y estrategias de búsqueda. A raíz de esta propuesta comienzan a aparecer en el mundo experiencias y propuestas de organización de los cursos de IA (Dillon, 1993).

Pero, los proyectos de inicios de los 80 fracasaron, varias empresas quiebran, lo cual condujo a una nueva crisis de la IA a inicios de los 90. Esta crisis no demoró mucho, en esa misma década resurgen nuevas investigaciones y resultados constantes en varios campos de

la IA. Las investigaciones se concentraron en un núcleo más definido y en resolver problemas más específicos del mundo real (Russell & Norvig, 2004).

A partir de la década del 90 se obtienen grandes resultados en los juegos. En 1994, un programa consigue ganar al campeón mundial de damas y se crea un programa para jugar backgammon utilizando aprendizaje por refuerzo y se introduce el algoritmo de Montecarlo para simular procesos estocásticos y así optimizar la elección de la jugada siguiente dentro del árbol de juego de Minimax. En 1997 la Deeper Blue, una máquina de carácter específico de IBM para jugar al ajedrez, logra derrotar a Kasparov, campeón del mundo de ajedrez.

En 2001 aparece una nueva propuesta de la ACM para currículos de Ciencias de la Computación (ACM-IEEE, Report, Chang, Lau, & Srimani, 2001).

### **Las principales diferencias entre la propuesta de 1991 y la de 2001 fueron (Urretavizcaya & Onaindía, 2002):**

- El incremento de la importancia en aspectos como la satisfacción de restricciones y la representación del conocimiento.
- Desaparece la utilización de los lenguajes LISP y Prolog como posibles lenguajes a utilizar en tareas prácticas.
- En los cursos complementarios y avanzados proponen los temas de razonamiento temporal, espacial, incertidumbre y conjuntos borrosos; desechando la representación basada en reglas.
- Aparece el tema de agentes, el cual ha venido a desempeñar un papel preponderante en la enseñanza de esta ciencia.

En esta etapa se evidencia un vertiginoso desarrollo de los currículos con una gran cantidad de temas tratados y diversas formas de enseñanza utilizadas (Bahamonde & Otero, 2001).

### **Algunos de los problemas planteados en la enseñanza de la IA en ese período fueron (Fasli, 2005):**

- Gran diversidad en los temas impartidos como básicos, por la falta de consenso entre los especialistas.
- Fuerte influencia de las experiencias prácticas en investigaciones y la especialización para decidir los contenidos de los cursos.
- Falta de metodologías y herramientas que puedan ser utilizadas para la enseñanza de estos temas.

En este período aparecen por primera vez, y se extienden hasta nuestros días, experiencias en la utilización de medios donde los estudiantes acceden al código fuente del software para incorporar los métodos y técnicas de IA abordados en las clases (Pantic, Zwitserloot y Grootjans, 2005) (Taylor, 2010) (Sosnowski, et. al. 2013) (Riedl, 2015) (Sintov, et. al. 2017). Dentro de las propuestas más interesantes se propone que cada profesor desarrolle su propia aplicación para llevar a cabo las prácticas de laboratorio (Kummeneje & Verhagen, 2005), de manera que integre todos los contenidos a tratar en clases (Mccluskey & Simpson, 2005). La gran mayoría han defendido la importancia de utilizar juegos como medios de enseñanza (Wang & Wu, 2009), aunque otros autores (Bellas & Alonso, 2007) defienden la importancia de abordar problemas reales o basados en la realidad, minimizando la utilización de códigos complejos.

Luego del 2010 vuelven a obtenerse resultados significativos en el área de los juegos. En febrero de 2011, el supercomputador Watson

de IBM gana un concurso televisivo contra jugadores humanos, en el que se realizan preguntas de todo tipo, cultura y conocimiento. Watson es capaz de aprender a medida que interactúa en lenguaje natural y acumula información. En el 2016, el software de inteligencia artificial Alpha Go de Google se enfrentó al campeón mundial de Go, un juego de estrategia complejo, en una partida a cinco juegos. Alpha Go ganó los tres primeros juegos limpiamente y solo en el último juego ganó el humano, gracias a un movimiento inicial que hizo y donde se comprobó que la máquina estaba poco entrenada para enfrentarse a situaciones inesperadas.

Un hecho de trascendencia histórica fue el curso de Introducción a la IA ofertado por Norvig y Thrun en 2011. Este curso fue montado en una plataforma en red y se ofertó de manera abierta para todo el mundo. Matricularon alrededor de 160 000 personas, siendo reconocido como el primer curso masivo y abierto en línea (MOOC por sus siglas en inglés) con éxito. A partir de ese momento se han ofertado varios cursos introductorios de IA en plataformas abiertas (Aznar et al., 2015), lo cual ha permitido ampliar el conocimiento de esta ciencia fuera del marco de las universidades.

En la siguiente propuesta de currículos (ACM-IEEE, 2013) para carreras de Ciencias de la Computación no se aprecian diferencias importantes. Los elementos de mayor interés son la profundización en el estudio de las estrategias básicas de búsqueda, la eliminación de los elementos de satisfacción de restricciones y la reducción del tiempo para la representación del conocimiento y formas de razonamiento. Se introduce por primera vez como parte del currículo básico el contenido de aprendizaje automático.

Otro tipo de aplicación de gran auge desde la segunda década de este siglo es la de los asistentes virtuales. Apple presenta el asistente virtual Siri integrado en el teléfono móvil iPhone en el año 2011. En

junio de 2012 Google presentó su asistente virtual, Google Now, y en abril de 2014 Microsoft presentó su propio asistente virtual, Cortana. Esta carrera se ha diversificado e intensificado en los últimos años.

En el propio 2014 un bot computacional llamado Eugene Goostman fue reconocido como el primero en superar la prueba de Turing al lograr que 30 de los 150 jueces asumieran que intercambiaban con un niño ucraniano de 13 años. Este resultado, aunque significativo, deja en evidencia la brecha que aún existe entre el comportamiento humano y el que se ha logrado llevar a los sistemas computacionales.

Aunque se aprecia relación entre los resultados prácticos que ha obtenido la IA y su enseñanza, quedan muchas brechas entre lo que se imparte en las carreras universitarias y lo que se trabaja en los centros de investigación y desarrollo. Un estudio llevado a cabo en Estados Unidos (Wollowski et al., 2016) encuestó a profesores y a investigadores de la IA para contrastar lo que proponían incluir en clases a los estudiantes. Como datos más significativos se destacan el deseo de los investigadores por lograr en los estudiantes habilidades en la ingeniería de sistemas, aspecto que no apareció en las respuestas de los educadores. Los educadores consideran importante tratar los temas desde una visión de juegos, mientras los investigadores desearían exponer a los estudiantes a sistemas con fuerte componente ingenieril con problemas del mundo real y de alta complejidad (Wollowski, Neller, & Boerkoel, 2017), (Sintov, Kar, Nguyen, & Fang, 2017). Los investigadores asumen como importante desarrollar habilidades en los estudiantes para utilizar técnicas y herramientas específicas de IA, mientras los docentes centran sus clases en elementos teóricos y generales.

Este estudio reafirma las siempre presentes contradicciones entre teoría y práctica en los diseños curriculares. Aunque las soluciones inmediatas estarán en buscar un justo equilibrio en las clases, se

puede trabajar en soluciones más integrales a mediano y largo plazo. Si se logra introducir elementos teóricos básicos e incluso prácticos sencillos desde la educación precedente, en la universidad se pueden dedicar mayores esfuerzos a los elementos prácticos y de mayor complejidad.

## La enseñanza de la IA en Cuba

En el año 1968 se reconocen los inicios de las investigaciones en IA en Cuba y en 1969 salen a la luz las primeras publicaciones para difundir conceptos de las ciencias de la computación y específicamente de la IA (MES-UH, 2018). En nuestro país, la enseñanza de la IA se inicia en el año 1977 con los planes de estudio A para carreras universitarias. En 1981 se propone una primera modificación denominada Plan B, en esta propuesta se propone en la carrera Cibernética-Matemática el estudio del Prolog y conceptos y métodos generales de la IA. En tanto, en la carrera Ingeniería Informática se introduce la lógica matemática y el Prolog.

En 1986 se elabora un nuevo conjunto de planes de estudio en Cuba (Plan C). En la carrera Ciencias de la Computación se profundiza en los elementos generales de la IA y se incluye la simulación y cursos optativos con temáticas específicas como las redes neuronales y el reconocimiento de patrones. En Ingeniería Informática, se enriquecen los contenidos con la asignatura Programación Descriptiva y Sistemas de Bases de Conocimientos I y II.

En 1998 se realiza una nueva revisión de los planes de estudio, dando paso a los planes C mejorados. En la carrera Ingeniería Informática se reduce el tiempo a la programación lógica y se renombran las asignaturas a Sistemas Informáticos Inteligentes I y Sistemas Informáticos Inteligentes II.

Desde finales de la década de los 90 se trabaja en utilizar medios en la enseñanza de la IA, sobre todo del Prolog, como lenguaje representativo del paradigma de programación lógica (Lezcano & Valdés, 1998) (Ríos, Lezcano, & Aljadis, 2008) (García & Lezcano, 2011) (Cruz, 2015).

Aunque no se han divulgado muchos estudios sobre la enseñanza de la IA en Cuba se plantean algunos elementos que han incidido en la enseñanza de esta ciencia (Lezcano, 2002).

### **Elementos que han incidido en la enseñanza de la IA (Lezcano, 2002):**

- Imprecisión en los objetivos de enseñanza: Contradicción entre el deseo de que los estudiantes aprendan un conjunto básico de conceptos o desarrollarles determinadas habilidades.
- La abstracción de los conceptos: Lo cual implica que sean difíciles de comprender por el estudiante.
- Imprecisión de las herramientas de medición del aprendizaje: No se logra saber qué saben los estudiantes, debido a que las herramientas tradicionales (exámenes, proyectos, etc.) no resultan suficientes para indicarlo.
- Contradicción entre el reducido tiempo que se le asigna en los planes de estudio y la diversidad y complejidad de los temas que trata.

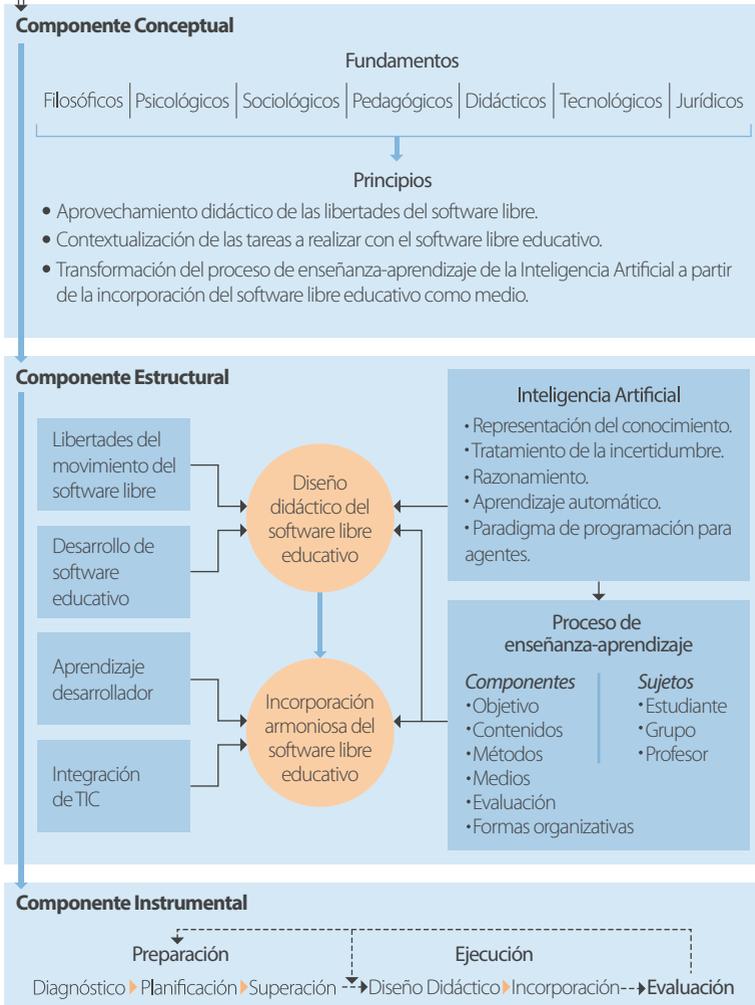
Estos elementos se mantienen o incluso se agravan con el desarrollo de la propia ciencia. Cada día aparecen nuevos algoritmos que incorporan más niveles de abstracción, se hace más complejo evaluar el aprendizaje de los estudiantes y se mantiene el dilema de conceptos contra habilidades, teoría contra práctica.

En el 2002 surge la Universidad de las Ciencias Informáticas (UCI) y con ella la carrera Ingeniería en Ciencias Informáticas (ICI). Su primera versión de plan de estudio incluyó una asignatura de IA con elementos básicos de búsquedas, representación del conocimiento, tratamiento de incertidumbre y fundamentos de los sistemas expertos.

En el 2008 se elabora un nuevo ciclo de planes de estudio en las carreras relacionadas con la informática en Cuba. Se abordan los fundamentos de la IA en general, de la teoría de decisión bajo incertidumbre y de los procesos de inferencia. Se incluyen modelos de representación y procesamiento del conocimiento, aprendizaje y minería de datos. En Ciencia de la Computación se abordan las formas de representación del conocimiento, estrategias de búsqueda y heurísticas asociadas en espacios de soluciones de problemas. Otros elementos incluidos son el aprendizaje automático y la aplicación de lenguajes declarativos, incluido el paradigma de programación lógica y el paradigma de programación funcional, para la solución de problemas. En la carrera Ingeniería en Ciencias Informáticas se incrementa a dos las asignaturas de IA. Como elementos más significativos se incorporan elementos de Prolog y aprendizaje automático.

Los más recientes planes de estudio en Cuba (Plan E), mantienen similares sistemas de conocimientos (MES-CUJAE, 2017) (MES-UH, 2018). Es significativo la integración lograda en Ingeniería Informática, donde los elementos de IA se integran con elementos de las matemáticas aplicadas, nombrando a la disciplina Inteligencia Computacional. Esta misma concepción se presentó en la carrera Ingeniería en Ciencias Informáticas. En las tres carreras de este perfil en Cuba, quedan asignaturas con el nombre histórico de IA para abordar los elementos de representación del conocimiento, estrategias de búsqueda y aprendizaje.

## Integración de software libre educativo al proceso de enseñanza-aprendizaje de la Inteligencia Artificial



**Legenda:** Subordinación → | Secuenciación - -> | Coordinación → | Entrada/Salida ==>

**Figura 3.** Modelo de integración de software libre educativo al proceso de enseñanza-aprendizaje de la Inteligencia Artificial.

Estudios más recientes sobre enseñanza de la IA en Cuba se dirigen a la integración de software libre educativo (Coca y Pérez, 2020) a las clases de inteligencia artificial (Coca y Pérez, 2021a). Para la integración de estas aplicaciones se propone un modelo que incluye desde el diseño didáctico de estas aplicaciones hasta su incorporación armoniosa al proceso de enseñanza-aprendizaje (Coca, 2021) (Figura 3). Como parte de estas investigaciones se han presentado estudios específicos con algunos resultados preliminares (Coca y Pérez, 2018) (Coca y Pérez, 2021b).

### **Futuro de la inteligencia artificial en el ámbito educativo**

El informe Horizon del 2019 pronostica que la IA en los próximos dos a tres años será importante para la enseñanza, el aprendizaje y la investigación creativa (Alexander et al., 2019). Exponen la preocupación con las cuestiones éticas, dado a que la IA desarrolla una capacidad cada vez más parecida a la de los humanos. El uso de los datos, la inclusión, el sesgo algorítmico y la vigilancia, se hacen cada vez más importantes. A pesar de estas preocupaciones, prevén un crecimiento significativo en las aplicaciones informáticas de la IA relacionadas con la enseñanza y el aprendizaje para la educación superior.

Dentro de los retos de la IA en el campo de la educación se destacan la comunicación en lenguaje natural, explicar sus razonamientos de alguna manera para ser más confiable para profesores y estudiantes y resolver problemas cada vez más parecidos a los que resuelve el hombre (Bello, 2019). Estos aspectos requieren seguir el desarrollo, tanto en el campo de la propia IA, como en los modelos pedagógicos que asuman estas tecnologías.

La comunicación en lenguaje natural generalmente incluye tanto el procesamiento de texto como el de voz. Esto requiere técnicas específicas de IA para cada sistema que se desarrolle. Sin embargo, hoy se suma la inclusión de imágenes, lo cual incorpora un nuevo reto, tanto en la visualización de avatares como en el análisis de videos. En todos los casos tomando en cuenta no solo los elementos del lenguaje, sino el contenido específico de la materia que se aborda. El elemento pedagógico requiere nuevas formas de presentar el contenido, dar seguimiento al trabajo del estudiante y evaluar su resultado, a partir de las correspondientes aplicaciones informáticas.

Explicar el razonamiento de las aplicaciones educativas basadas en Inteligencia Artificial para ser más confiables para profesores y estudiantes, es un gran reto. Varias técnicas de Inteligencia Artificial tienen un funcionamiento interno matemático y complejo imposible de explicar en lenguaje natural a los profesores y estudiantes. Esto muchas veces hace desconfiar de los resultados, aunque estadísticamente los resultados sean muy favorables. Por otro lado, las varias fuentes de incertidumbre, aunque tratadas desde la propia IA, brindan resultados no deterministas.

Aunque todo el desarrollo de la IA ha estado, y sigue siendo, en la dirección de resolver problemas cada vez más parecidos a los que resuelve el hombre, aún es muy distante en algunos campos. En el ámbito educativo, el profesor sigue siendo la figura imprescindible en el proceso de enseñanza-aprendizaje. Tardará mucho, o nunca llegará, el momento en que el aprendizaje sea enteramente con aplicaciones informáticas. La formación en valores, las cuestiones éticas, los procesos psicológicos, entre muchos otros, son grandes retos de la IA para aplicaciones educativas.

Aunque la enseñanza de esta ciencia requiere estar a tono con estos retos, exige buscar las invariantes en el sistema de conocimientos

a brindar a los estudiantes. Se hace necesario acercar las asignaturas a los problemas profesionales que el futuro egresado deberá enfrentar en la economía y la sociedad en su conjunto. La formación integral para resolver problemas de IA desde los más simples hasta los más complejos, requiere identificar ese sistema de conocimientos básico y las vías más efectivas de desarrollar las habilidades necesarias en el futuro profesional.

En los últimos años el desarrollo de la robótica, incluida la conducción autónoma de autos genera nuevas noticias casi a diario. El aprendizaje profundo, con el desarrollo de nuevos modelos de redes neuronales y otras técnicas, ha venido a desarrollar aplicaciones que logran grandes resultados a partir del gran cúmulo de información que se genera en nuestros días.

Las investigaciones en IA han ido avanzando más rápidamente, gracias a un mayor uso del método científico en la experimentación y comparación de propuestas. Los avances recientes logrados en el entendimiento de las bases teóricas de la inteligencia artificial han ido aparejados con las mejoras realizadas en la optimización de los sistemas reales. La IA ha encontrado elementos comunes con otras disciplinas y sus campos de estudio se han integrado más.

El paradigma científico, es el proceso de formar hipótesis y verificarlas a través de experimentos. Las hipótesis con éxito son vistas como modelos que explican y predicen fenómenos en el mundo. La Computación sigue este paradigma y se reconoce como su objeto de estudio el procesamiento de la información (Denning, 2005). Aunque algunos detractores la incluyen como disciplina dentro de las Matemáticas o el Control. Otros la asumen como tecnología o arte, pero no como Ciencia.

La Ciencia de la Computación estudia el procesamiento de información en procesos que ocurren naturalmente en el mundo fí-

sico. Los científicos de la computación trabajan con un aceptado y sistemático cuerpo de conocimiento. Gran parte de la ciencia de la computación es ciencia aplicada y es utilizada para predicción y verificación. La IA, como parte de las Ciencias de la Computación, se abre camino como ciencia independiente con su propio objeto de estudio y problemas científicos por resolver.

En nuestro país se han llevado a cabo investigaciones que plantean rasgos esenciales de la ciencia en sentido general y que permitirían determinar si una rama específica es ciencia (Páez, 2013). Algunos de esos rasgos pudieran abordarse con mayor profundidad, como su carácter humanista y su respuesta a las necesidades del proyecto social y la política estatal. Otros, brindan fuertes criterios para asumir que la IA se abre paso como una Ciencia de increíble desarrollo e impacto en toda la sociedad.

### **Rasgos que permitirían determinar si una rama específica puede ser considerada ciencia (Páez, 2013):**

- Sistema de conocimientos teóricos que modifica nuestra visión del mundo.
- Proceso investigativo de producción de conocimientos que enriquece la cultura.
- Proceso que deviene fuerza productiva que propicia la transformación del mundo.
- Tiene un carácter humanista y tecnológico.
- Es a la vez institución social, en un sistema de organizaciones científicas.
- Responde a las necesidades del proyecto social y de la política estatal.

Aunque algunos de los puntos pudieran abordarse con mayor profundidad, como su carácter humanista y su respuesta a las necesidades del proyecto social y la política estatal, se puede afirmar que la Inteligencia Artificial responde a estos aspectos y por tanto se abre paso como una Ciencia de increíble desarrollo e impacto en toda la sociedad.



## LA NECESARIA REFLEXIÓN ÉTICA

---

La IA encierra el potencial de mejorar el futuro de la humanidad y promover el desarrollo sostenible. Sin embargo, va creciendo la conciencia en el mundo de los riesgos y desafíos asociados con la misma, especialmente en lo que se refiere al incremento de las desigualdades y brechas existentes, así como las implicaciones para los derechos humanos. Aunque ya se han mencionado algunos aspectos éticos en el ámbito educativo, en este capítulo abordaremos con mayor profundidad los riesgos, amenazas y el trabajo que se lleva a cabo para lograr un desarrollo ético y sostenible de la IA.

Algunas aplicaciones en nuestros días ya son cuestionables, por ejemplo, la recogida de datos que invaden la vida privada, algoritmos de reconocimiento facial que deben identificar conductas hostiles o están imbuidos de prejuicios raciales, drones militares y armas letales autónomas, etc (Editorial, 2018).

En las redes digitales y el internet de forma general, la información sobre nuestras preferencias y hábitos es utilizada para crear modelos de comportamiento. Dichos modelos permiten que los sistemas de IA envíen mensajes de carácter político, ofrezcan aplicaciones comerciales o, incluso, almacenen información relativa a nuestros cuidados médicos (Wee Ang & Feinholz, 2018). Además,

la automatización y la digitalización crean nuevos desequilibrios, pueden reducir la diversidad en las industrias culturales, transforman el mercado de trabajo, generan precariedad y aumentan las desigualdades entre los que tienen acceso a estas nuevas tecnologías y los que no.

*Al igual que otras muchas tecnologías de la información, la IA plantea cuestiones relacionadas con la libertad de expresión, la privacidad, la vigilancia, la propiedad de los datos, el sesgo, la discriminación, manipulación de la información, la confianza, las relaciones de poder, el impacto medioambiental, entre muchas otras. Además, la IA plantea específicamente nuevos retos relacionados con su interacción con las capacidades cognitivas humanas.*

Los sistemas basados en la IA tienen consecuencias para la comprensión y la experiencia humana. Las redes digitales y los sitios de noticias pueden facilitar la propagación de desinformación, así como la interacción y la participación en el ámbito político. Los algoritmos diseñados para evitar el sesgo político humano a la hora de decidir qué contenido aparecerá en las redes, pueden utilizarse para promover deliberadamente la difusión de contenido falso, manipulador y generador de división, dirigido a grupos objetivo específicos. En algunos casos, este contenido puede incluir información formateada de manera fraudulenta como noticias, así como elementos que sirvan como propaganda emocional.

Por otro lado, el desarrollo de la IA acentúa la asimetría entre individuos, grupos y países, incrementan la brecha digital en cada nación, y entre naciones. Tal brecha puede verse agravada por la falta de acceso a elementos fundamentales como los algoritmos para el aprendizaje y la clasificación, los datos para entrenar y evaluar los

algoritmos, los recursos humanos para codificar y configurar el software y preparar los datos, así como los recursos computacionales para el almacenamiento y el procesamiento de los datos.

Muchos sistemas ya hoy traspasan nuestras facultades cognitivas en diversos terrenos, lo cual despierta temores de riesgos de carácter ético. Estos riesgos pueden ser de tres órdenes (Ganascia, 2018): la escasez de trabajo, que sería ejecutado por máquinas en lugar de seres humanos; las consecuencias para la autonomía del individuo, en especial para su libertad y su seguridad; y la superación del género humano, que sería sustituido por máquinas cada vez más inteligentes. Sin embargo, se ha demostrado que el trabajo no desaparece, sino que, muy al contrario, se transforma y exige nuevas habilidades. La autonomía del individuo y su libertad no están comprometidas por el desarrollo de la IA, siempre y cuando nos mantengamos alerta en lo relativo a las intromisiones de la tecnología en nuestra vida privada.

Las máquinas de forma general no constituyen una amenaza existencial para la humanidad. Su autonomía es de carácter técnico, es decir, se encargan de los procesos que van desde la búsqueda de información hasta la toma de decisiones. Pero, no tienen autonomía moral, no poseen voluntad propia y permanecen al servicio de los objetivos que se les han fijado. Aunque, cada vez más el propio desarrollo de la IA se encargue de hacernos ver que son verdaderamente inteligentes y conscientes de lo que hacen.

Uno de los aspectos más peligrosos que hay que atender con el desarrollo de la IA es el jurídico y militar. La investigación en materia de IA avanza a gran velocidad, mientras que los contextos jurídicos, sociales y éticos en los que debería enmarcarse evolucionan muy lentamente. Ya son numerosas las aplicaciones en la industria militar, incluso proporcionando aliados robots a los combatientes. Facilita

maniobras y mejora el rendimiento de los ejércitos. Según algunos expertos, los Sistemas Armamentísticos Autónomos Letales (SALA, por sus siglas en inglés) están creando la tercera revolución bélica, después de la pólvora y las armas nucleares. Se espera que pronto llegará el día en que los ejércitos de robots serán capaces de realizar hostilidades con total autonomía (Sychev, 2018).

En la actualidad, muchos empresarios, responsables políticos y científicos tratan de prohibir el uso de sistemas de combate autónomos, mientras que los militares aseguran que en el combate la decisión final será tomada por un ser humano. Esto no reduce la amenaza, ya que nadie puede asegurar que, como ocurrió con el arma nuclear, sea utilizada a pesar de haberse enfrentado a numerosos opositores (Sychev, 2018).

Los sistemas de IA tienen implicaciones significativas para la igualdad de género, ya que pueden reflejar los sesgos sociales existentes y hasta fomentarlos. La mayoría de los sistemas de IA se construyen utilizando conjuntos de datos que reflejan el mundo real, que puede ser imperfecto, injusto y discriminatorio (Marda, 2018). Los algoritmos para clasificación, por ejemplo, utilizan datos específicos de una cultura y son producto de una visión específica de la historia, por lo que pueden generar un sesgo en las decisiones adoptadas. Si el mismo algoritmo utilizara diversos conjuntos de datos, su sesgo puede reducirse, pero sería computacionalmente imposible suprimirse por completo (Conferencia General, 2019).

El desarrollo de la IA debe ir acompañado de una reflexión ética, ya que las tecnologías de IA no son neutrales, sino que están intrínsecamente sesgadas por los datos en los que se basan y las decisiones que se toman durante la integración de esos datos. A la larga, no tan larga, las decisiones que tomarán las máquinas inteligentes no podrán predecirse o controlarse totalmente.

## El debate sobre ética

Para que la IA pueda trabajar a nuestro servicio, y no a nuestra costa (Wee Ang & Feinholz, 2018), debe propiciarse un debate de fondo que tenga en cuenta los puntos de vista éticos de todos los implicados. Debe vigilarse el marco ético en el que se enmarca su desarrollo futuro, teniendo en cuenta la responsabilidad social que encierra.

Desde hace algunos años se han encaminado esfuerzos para propiciar un debate mundial sobre ética en el desarrollo de la IA. En 2014, la ONU inició un debate internacional sobre el desarrollo de los SALA. En 2017 un Grupo de Expertos Gubernamentales (GEG), recibió el encargo de examinar las tecnologías emergentes en materia de SALA. Una de las orientaciones establecidas en el informe de consenso de esta reunión es que la responsabilidad por el desarrollo de cualquier sistema de armamento de conflicto corresponde a los Estados. Se destacó la necesidad de un entendimiento común con respecto a los posibles límites del grado de autonomía en el uso de la fuerza letal (Sychev, 2018).

La Universidad de Montreal, en 2017, comenzó un debate con miras a la elaboración de la Declaración de Montreal para un desarrollo responsable de la IA. Esta iniciativa tiene como principal objetivo establecer orientaciones éticas en el desarrollo de la IA a escala nacional. En debates entre el público en general, expertos y responsables políticos, se identificaron siete valores: bienestar, autonomía, justicia, vida privada, conocimiento, democracia y responsabilidad (Bengio, 2018).

La UNESCO por su parte ha liderado un debate para definir las dimensiones éticas y mundiales de la paz, la diversidad cultural, la igualdad de género y la sostenibilidad, en relación con su labor. De

forma general brindan criterios de utilidad para un desarrollo ético y sostenible de la IA. El Grupo de Trabajo de la UNESCO propone una serie de principios genéricos para el desarrollo, la implantación y el uso de la IA (Conferencia General, 2019) (Figura 4).

- a) **Derechos humanos:** debe desarrollarse e implementarse de acuerdo con las normas internacionales de los derechos humanos.
- b) **Integración:** debe ser inclusiva, con el objetivo de evitar sesgos, propiciar la diversidad y prevenir una nueva brecha digital.
- c) **Prosperidad:** debe desarrollarse para mejorar la calidad de vida.
- d) **Autonomía:** debe respetar la autonomía humana mediante la exigencia del control humano en todo momento.
- e) **Explicabilidad:** debe ser explicable, capaz de proporcionar una idea de su funcionamiento.
- f) **Transparencia:** los datos utilizados para entrenar los sistemas deben ser transparentes.
- g) **Conocimiento y capacitación:** se hace necesario, para capacitar a los ciudadanos, el conocimiento de los algoritmos y una comprensión básica del funcionamiento de los sistemas.
- h) **Responsabilidad:** los desarrolladores y las empresas deben tener en cuenta la ética al desarrollar los sistemas inteligentes autónomos.
- i) **Asunción de responsabilidades:** Deben desarrollarse mecanismos que permitan atribuir responsabilidades respecto a las decisiones basadas en la IA y el comportamiento de los sistemas.
- j) **Democracia:** debe desarrollarse, implantarse y utilizarse con arreglo a principios democráticos.
- k) **Buena gobernanza:** los gobiernos deben presentar informes periódicos sobre su utilización en los ámbitos de la actividad policial, la inteligencia y la seguridad.

**1) Sostenibilidad:** En todas las aplicaciones, los beneficios potenciales deben equilibrarse con el impacto medioambiental del ciclo de producción completo de la IA y las tecnologías de la información.



**Figura 4.** Principios éticos de la IA.

## En busca de consensos

La UNESCO presenta algunas cuestiones éticas fundamentales en relación con el enfoque específico de la organización. Las cuestiones éticas más importantes en lo que concierne a la IA atañen a sus implicaciones para la cultura y la diversidad cultural, la educación, el conocimiento científico, la comunicación y la información. Además, dada la orientación mundial de la UNESCO, los temas éticos

globales de la paz, la sostenibilidad, la igualdad de género, y los retos específicos para África también merecen una atención específica (Conferencia General, 2019) (Figura 5).

La IA exige un replanteamiento del papel social de la **educación**. El desplazamiento laboral causado por algunos sistemas de IA requiere, entre otras medidas, la recapacitación profesional de los trabajadores y un nuevo enfoque para evaluar los programas educativos.

Aunque ya está ocurriendo, se prevé que se profundice la transformación de la educación, la modificación de los métodos de enseñanza, las formas de aprender, de acceder al conocimiento, de capacitar, no solo a los estudiantes, también a los docentes. Las competencias que habrá que desarrollar para evolucionar en un mundo cada vez más automatizado será un tema que adquirirá cada vez mayor importancia. La educación deberá habilitar a los ciudadanos para que desarrollen nuevas formas de pensamiento crítico, incluido el análisis y crítica de los algoritmos, la capacidad de reflexionar sobre el impacto de la IA en la información, el conocimiento y la toma de decisiones.

### Educación

Fomentar la adquisición de competencias en materia de IA, el pensamiento crítico, la resiliencia en el mercado laboral y la instrucción ética de los ingenieros.

### Ciencia

Introducirla de forma responsable en la práctica científica y en la toma de decisiones, exigiendo la evaluación y el control humanos y que se evite la exacerbación de las desigualdades estructurales.

### Cultura

Fomentar la diversidad cultural, la inclusión y la experiencia humana, evitando una profundización de la brecha digital. Promover un enfoque multilingüe.

### Comunicación e información

Consolidar la libertad de expresión, el acceso universal a la información, la calidad del periodismo y los medios libres, independientes y pluralistas, evitando la propaación de la desinformación.

Paz	Obtener información sobre factores que impulsan los conflictos y nunca funcionar al margen del control humano.
África	Integrarla en las políticas y estrategias de desarrollo nacional sobre la base de las culturas, los valores y los conocimientos endógenos para desarrollar las economías africanas.
Igualdad de género	Evitar el sesgo de género en el desarrollo de algoritmos, en los conjuntos de datos utilizados para su formación, y en su uso en la toma de decisiones.
Medio ambiente	Desarrollarla de manera sostenible, teniendo en cuenta todo el ciclo de producción de la IA y las TI. Utilizarla en la vigilancia medioambiental, gestionar riesgos, prevenir y mitigar las crisis medioambientales.

**Figura 5.** Cuestiones éticas fundamentales de la IA.

Otro ámbito de las cuestiones éticas es la relativa al papel de la IA en el proceso educativo en sí, su amplia utilización en los entornos virtuales de enseñanza-aprendizaje, la robótica educativa y los sistemas de análisis del aprendizaje, todos requieren un desarrollo y una implementación responsables. Los desarrolladores de software deben recibir la formación adecuada para garantizar el diseño y la implantación responsables de la IA.

Hasta la fecha, los modelos educativos se han centrado en proporcionar conocimientos básicos, competencias formales como la lectura, la escritura y las matemáticas. En nuestros días, la información y el conocimiento están disponibles, lo cual exige una capacitación en materia de datos e informaciones que permita a los alumnos leer, analizar y gestionar eficazmente esta información, pero, también una capacitación que permita la reflexión crítica sobre el modo en que los sistemas informáticos inteligentes han participado en el reconocimiento de las necesidades de información, en la selección, la interpretación, el almacenamiento, la representación del conoci-

miento, el razonamiento y el tratamiento de la incertidumbre que realiza.

El pensamiento crítico será una de las primeras competencias que tendrán que transmitir los sistemas educativos, pero no la única. Las competencias esenciales pueden adquirirse en actividades extraescolares donde se aprenden, por ejemplo, lo que son la cooperación, el establecimiento de objetivos, la planificación.

*Se desarrolla la disciplina y el espíritu de equipo haciendo deporte; la creatividad, haciendo teatro; el pensamiento crítico, organizando debates; y la empatía, haciendo tareas sociales y extensionistas. Pero desde la IA existen muchas otras posibilidades. Se desarrolla la disciplina y el espíritu de equipo trabajando juntos por un bien común; la creatividad, desarrollando o adaptando nuevos algoritmos; el pensamiento crítico, definiendo heurísticas, seleccionando datos o algoritmos específicos para una tarea; y la empatía, brindando y compartiendo nuestros resultados como bibliotecas o software libre.*

El desafío consiste en estructurar este gran abanico de competencias que los jóvenes deben adquirir. Hay que oficializarlos en los sistemas educativos e integrarlos a los programas escolares y definir la manera de evaluar los resultados de los alumnos en estos dominios, no considerados hasta ahora como parte de la educación escolar.

Si un país dispone ya de una infraestructura adecuada y de un sistema educativo de buena calidad, puede sacar provecho de la IA para mejorar la eficacia de su producción, de la misma manera que el uso de la máquina a vapor permitió que algunos países se desarrollaran más rápido que otros durante la revolución industrial (Qiang, 2018).

Tradicionalmente, los ingenieros se han formado para desarrollar productos que optimicen el rendimiento, para lo cual intentan minimizar recursos, a partir de determinadas restricciones dadas. La ética de la tecnología ha desarrollado diversos métodos para incorporar la reflexión crítica, la responsabilidad y el razonamiento a este proceso. En el contexto de la IA, se ha acuñado el término Diseño Éticamente Alineado (DEA) para referirse a los procesos de diseño en los que se consideran explícitamente valores humanos.

La formación en ingeniería a escala global se centra fundamentalmente en cursos científicos y tecnológicos que no están intrínsecamente relacionados con los análisis de los valores humanos diseñados abiertamente para reforzar positivamente el bienestar humano y medioambiental. Resulta esencial cambiar esta situación e instruir a los futuros ingenieros y científicos informáticos para que adopten el diseño éticamente alineado de los sistemas de IA. Esto requiere una concienciación explícita de las posibles implicaciones sociales y éticas, y de las consecuencias de la tecnología en el diseño, así como de su posible uso indebido.

La IA necesita introducirse de forma responsable en la práctica científica y en la toma de decisiones. La validez de una decisión basada en la IA debe tratarse con precaución. Tal decisión no es necesariamente equitativa, justa, precisa o apropiada. Es susceptible de inexactitudes, resultados discriminatorios, sesgos implícitos o incorporados, y limitaciones del proceso de aprendizaje. Un ser humano no solo posee una visión del mundo mucho más amplia, sino que también cuenta con un conocimiento tácito que supera a los sistemas de IA en situaciones críticas y complejas. Idealmente, la decisión de un sistema de IA sería similar a la de un humano si hubiera podido procesar una inmensa cantidad de datos de diversos dominios en un plazo razonablemente corto. Incluso así, los seres

humanos poseen capacidades diferentes y toman decisiones basadas en otros aspectos como la sensibilidad, la justicia y la ética.

La calidad de los sistemas de IA depende, en gran medida, de los datos disponibles utilizados para entrenar a los algoritmos. Muchos de los sistemas más importantes hoy son desarrollados por empresas privadas, lo cual implica que no siempre existe suficiente transparencia respecto a los datos utilizados. Este aspecto contrasta con el método científico tradicional que garantiza la validez de los resultados al exigir la replicabilidad, es decir, la posibilidad de reproducirlos mediante la repetición de los mismos experimentos.

En otro aspecto importante, aunque la IA es una poderosa herramienta para la creatividad humana, plantea importantes interrogantes sobre el futuro del arte y los derechos de los artistas. La IA debe fomentar la **diversidad cultural**, la inclusividad y el florecimiento de la experiencia humana, procurando no ampliar la brecha digital. También deberá promoverse una estrategia plurilingüe.

La IA se emplea con frecuencia en el ámbito de la cultura, por ejemplo, en las imágenes 3D utilizadas para la reconstrucción del patrimonio. Abordar el asunto de la diversidad cultural comprende la tarea de garantizar un esfuerzo activo encaminado a la inclusión de género, así como al fomento de la diversidad social y cultural de los ingenieros, a una consideración global de las implicaciones sociales y éticas del diseño de los sistemas de IA.

La IA desempeña un papel cada vez más importante en el procesamiento, estructuración y suministro de **información**. Resulta indispensable prestar atención a las nuevas brechas digitales entre países y dentro de distintos grupos sociales. Se debe trabajar por fortalecer la verdadera libertad de expresión, el acceso universal a la información y la calidad del periodismo, sin sesgos políticos, culturales o de cualquier otra índole. En imprescindible lograr que

los sistemas y algoritmos de IA no brinden beneficios a los grandes medios hegemónicos y logren ponderar la importancia de medios de comunicación con visiones diferentes, independientes, pluralistas, pero, sobre todo, evitando la difusión de información falsa. Será importante el fomento de una gobernanza de Internet en la que participen múltiples partes interesadas.

Con el fin de contribuir a la **paz**, la IA podría utilizarse para obtener información sobre los factores que impulsan los conflictos, y brindar opciones para resolverlos. La nueva economía digital que está surgiendo trae consigo importantes desafíos y oportunidades para las sociedades de África y otros países en desarrollo. El trabajo debe ser intencionado para revertir el incremento de la brecha digital que se aprecia en nuestros días. Desde el punto de vista ético, la IA deberá integrarse en las políticas y estrategias nacionales de desarrollo, dando cabida a las culturas, valores y conocimientos endógenos a fin de desarrollar las economías nacionales y locales.

La IA es una tecnología extendida y, en teoría, accesible a todos. Sin embargo, la gobernanza actual se ejerce por un grupo de instituciones, organizaciones y empresas. La reflexión sobre su **buena gobernanza** requiere un enfoque más pluralista, multidisciplinario, multicultural y multipartito, que plantee interrogantes sobre el tipo de futuro que queremos para la humanidad. Esta reflexión debe abordar los principales desafíos derivados del desarrollo de tecnologías de IA en cuanto a los sesgos incorporados en los algoritmos, incluidos los prejuicios de género, la protección de la privacidad de las personas y los datos personales, los riesgos de crear nuevas formas de exclusión y desigualdad, las cuestiones vinculadas a la distribución justa de los beneficios y los riesgos, la rendición de cuentas, la responsabilidad, las consecuencias para el empleo y el futuro del trabajo, la dignidad y los derechos humanos.

Deberá evitarse el sesgo de **género** en la elaboración de algoritmos, en los conjuntos de datos utilizados para su creación, y en su uso para la toma de decisiones. La IA deberá desarrollarse de manera sostenible, teniendo en cuenta el ciclo completo de su producción y de las tecnologías de la información. De igual forma, la IA puede emplearse para la vigilancia ambiental y la gestión de riesgos, y para prevenir y atenuar las crisis ambientales.

De forma general se puede afirmar que la inteligencia artificial tiene importantes implicaciones sociales y culturales. Puede generar predicciones muy precisas, basadas en conjuntos de datos, sin proporcionar ninguna explicación causal o unificadora de sus predicciones. Sus algoritmos no funcionan con los mismos conceptos semánticos que emplean los seres humanos para alcanzar la comprensión científica de un fenómeno. Esta brecha entre las predicciones de éxito, por un lado, y un entendimiento científico satisfactorio, por el otro, desempeñará probablemente un papel clave en la práctica científica, así como en la toma de decisiones basada en la IA.

*Un gran reto en nuestros días es lograr algoritmos explicables (Bello, 2019) cuando predominan algunos como las redes neuronales o las máquinas de soporte vectorial con un procesamiento matemático oculto que desafía el pensamiento humano. Este hecho podría repercutir en la confianza en la ciencia, que se basa habitualmente en el método científico que explica los diferentes fenómenos de una manera sistemática y transparente, realizando predicciones racionales y basadas en datos contrastados.*

Algunas de las amenazas atribuidas a la IA son improbables, pero merece la pena analizarlas todas y desarrollar acciones para que efectivamente no se hagan realidad. Queda claro que el desarrollo

de la IA puede conducir a máquinas inteligentes que conducirían a un futuro incierto, y que puede no ser de nuestro agrado. Se hace evidente la posibilidad de que individuos puedan emplear armas, sistemas de IA o robots que impliquen destrucción masiva y el fin de la propia humanidad. Pero, estas amenazas están presentes en muchas otras ciencias y se hace, incluso, más probable en ramas como la biotecnología o la nanotecnología que en la propia robótica (Russell & Norvig, 2004).



---

## CONSIDERACIONES FINALES

---

La Inteligencia Artificial es una Ciencia relativamente joven. Su desarrollo ha estado marcado por etapas de gran auge y otras de estancamiento, sobre todo por previsiones muy optimistas que la han acompañado y que no se han podido cumplir en el tiempo. No se puede negar que en el corto tiempo que lleva de desarrollo consciente, ha alcanzado avances muy significativos. En nuestros días es motor de muchas otras ciencias e impacta gran cantidad de áreas de la economía y la sociedad en su conjunto.

Como toda Ciencia, la IA debe ir acompañada de un componente ético en su aplicación. Por los temas que aborda para la toma de decisiones y el comportamiento de sistemas de forma autónoma, exige un debate, trazar pautas y encaminar acciones para su desarrollo sostenible en consonancia con el desarrollo de la propia humanidad.





## GLOSARIO DE TÉRMINOS

---

**Acciones (de un agente):** Evidencias del comportamiento del agente en su entorno.

**Agente (racional):** Entidad computacional que se comporta a partir de maximizar una medida de rendimiento cuyo valor depende de las percepciones realizadas de su entorno.

**Aplicaciones (de la IA):** Áreas donde se evidencia la utilización práctica de los conocimientos sobre una materia, relacionándose con otras materias científicas o ingenieriles. Las aplicaciones de la IA incluyen otras ramas de las ciencias de la computación, otras ciencias y sectores específicos de la economía y la sociedad. Resaltan en nuestros días el procesamiento de lenguaje natural, la minería de datos, los videojuegos y la robótica.

**Aprendizaje (en un agente):** Utilizar las percepciones de su entorno para comportarse y mejorar su habilidad para actuar en el futuro.

**Aprendizaje no supervisado:** El sistema se comporta a partir de patrones de entradas para los que no se especifican los valores de sus salidas.

**Aprendizaje reforzado:** El sistema actúa a partir de la información que obtiene de su propio comportamiento en el entorno.

**Aprendizaje supervisado:** El sistema se comporta a partir de la utilización de ejemplos con sus entradas y sus salidas.

**Creencia (de un agente):** Representan el conocimiento que se tiene del entorno. Desde un punto de vista informático, son la forma de representar el estado del entorno.

**Comportamiento racional (de un agente):** Función del agente que proyecta una percepción dada, en una acción.

**Deseos (de un agente):** Son los objetivos del agente. Un objetivo se asocia al valor de una variable, un registro, o una expresión simbólica en alguna lógica. Representa un estado final deseado.

**Entorno (de un sistema computacional):** Conjunto de datos, informaciones, conocimiento, circunstancias, factores, etc., que influyen en el estado o desarrollo de un sistema computacional.

**Heurística:** Criterio que evalúa la posibilidad de que una búsqueda vaya en la dirección correcta.

**Incertidumbre (en IA):** Falta de conocimiento sobre el entorno.

**Inteligencia Artificial:** Es la Ciencia de la Computación encargada de aplicar métodos de representación del conocimiento, razonamiento, tratamiento de la incertidumbre y aprendizaje, en el desarrollo de sistemas informáticos con comportamiento racional.

**Intenciones (de un agente):** Conjunto de planes especificados por el agente para alcanzar los objetivos propuestos, a partir de las creencias existentes.

**Medida de rendimiento (de un agente):** Criterio utilizado para valorar el desempeño del agente en su entorno.

**Minería de datos:** Aplicación práctica de la Inteligencia Artificial.

Usualmente se asocia al aprendizaje automático, aunque realmente requiere de los cuatro núcleos de conocimiento de la IA.

**Núcleo de conocimientos (de la IA):** Temáticas básicas en las que se estructura una materia o disciplina para su estudio. Los núcleos, estrechamente relacionados entre sí, de la IA son: la Representación del Conocimiento, el Razonamiento, el Tratamiento de la Incertidumbre y el Aprendizaje.

**Percepción (de un agente):** Son las entradas que recibe el agente en cualquier instante. La secuencia de percepciones de un agente refleja el historial completo de lo que el agente ha recibido.

**Procesamiento de lenguaje natural:** Aplicación de la IA encargada de analizar texto o voz utilizando un conjunto de teorías y tecnologías. Incluye la comprensión del lenguaje y la generación del lenguaje. Aunque algunos lo consideran un núcleo de conocimiento de la IA, lo cierto es que requiere de los cuatro núcleos de conocimiento de la IA y otras áreas para ser abordado.

**Prueba de Turing:** Un sistema supera la prueba si un evaluador humano no es capaz de distinguir si las respuestas, a una serie de preguntas planteadas, son de una persona o del sistema.

**Racionalidad:** Se considera que un sistema o agente es racional si hace lo correcto, en función de su conocimiento.

**Razonamiento (computacional):** Proceso general llevado a cabo por un sistema para comportarse de manera racional a partir del conocimiento que tenga de su entorno.

**Red neuronal artificial:** Modelo matemático desarrollado en el área de la Inteligencia Artificial que trata de representar el funcionamiento del cerebro. Se utiliza para tareas de clasificación y

regresión fundamentalmente, incluye un proceso previo de entrenamiento. Ha acompañado los principales resultados de la IA desde sus inicios hasta nuestros días, por lo que es una de las técnicas más reconocidas en este campo.

**Representación del conocimiento:** Proceso de estructurar el conocimiento mediante una notación suficientemente precisa como para ser utilizado por el sistema en su comportamiento.

**Robótica:** Área de aplicación de la IA con mayor complejidad. Integra otras aplicaciones directas de la IA como la visión por computador y el procesamiento de lenguaje natural, así como acciones complejas como el movimiento, equilibrio, tacto, entre otras. Requiere integración con áreas como la mecánica, la electrónica y muchas otras ciencias e ingenierías.

**Sistema multiagente:** Sistema computacional donde varios agentes actúan en el mismo entorno para dar cumplimiento a tareas específicas.

**Utilidad (de un agente):** Medida de rendimiento que indica lo deseables que son los estados posibles del agente y las acciones que va a ejecutar para llegar a ellos.

**Visión por computador:** Aplicación de la IA que analiza imágenes o secuencia de imágenes para tomar decisiones sobre ellas. Requiere integrar otras áreas como los gráficos por computador y la física, en muchos casos además, la biología e incluso la psicología.

---

## REFERENCIAS

---

- ACM-IEEE-CS. (1991).** COMPUTING CURRICULA 1991. Communications of the ACM, 34(6), 69–84.
- ACM-IEEE. (2013).** CS2013: Computer Science Curricula 2013. En Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. <https://doi.org/10.1145/2534860>
- ACM-IEEE, Report, F., Chang, C., Lau, F., y Srimani, P. (2001).** Computing Curricula 2001. Computer Science.
- ACM. (2020).** Association for Computing Machinery. Advancing Computing as a Science & Profession. Accedido desde el sitio oficial: <https://www.acm.org/>
- Alexander, B., Ashford-Rowe, K., Barajas-Murphy, N., Dobbin, G., Knott, J., McCormack, M., ... Weber, N. (2019).** EDUCAUSE Horizon Report: 2019 Higher Education Edition. <https://library.educause.edu/-/media/files/library/2019/4/2019horizonreport.pdf?la=en&hash=C8E8D444AF372E705FA1B-F9D4FF0DD4CC6F0FDD1>
- Atchison, W. F., Schweppe, E. J., Viavant, W., Young, D. M., Conte, S. D., y Hamblen, J. W. Rheinboldt, W. C. (1968).** Curriculum 68: Recommendations for academic programs in computer science: a report of the ACM curriculum committee

- on computer science. *Communications of the ACM*, 11(3), 151–197. <https://doi.org/doi:10.1145/362929.362976>
- Austing, R. H., Barnes, B. H., Bonnette, D. T., Engel, G. L., y Stokes, G. (1979).** Curriculum '78: recommendations for the undergraduate program in computer science. *Communications of the ACM*, 22(3), 147–166.
- Aznar, F., Compañ, P., Pujol, M., Rizo, R., Rizo, C., Sempere, M., ... Viejo, D. (2015).** Consideraciones iniciales del cambio de metodología presencial a MOOC. El caso de Sistemas Inteligentes. *Jornadas de Redes de Investigación En Docencia Universitaria*, 1465–1475. Universidad de Alicante.
- Bahamonde, A., y Otero, R. (2001).** ESPECIAL CAEPIA-TTIA ' 2001. IX Conferencia de La Asociación Española Para La Inteligencia Artificial. Gijón, España.
- Bellas, F., y Alonso, A. (2007).** Metodología de trabajo y experiencias de aprendizaje colaborativo y evaluación continua en la disciplina de Sistemas Multiagente. XIII Jornadas de Enseñanza Universitaria de La Informática, 65–73.
- Bello, R. (2019).** Trends of Artificial Intelligence: challenge in Education. *International Conference on Interactive Collaborative and Blended Learning*, 51. Santiago de Cuba.
- Bengio, Y. (2018).** Resistir a la monopolización de la investigación. *El Correo de La UNESCO*, 2018(3), 18–19.
- Coca, Y. (2021)** Modelo de Integración de software libre educativo al proceso de enseñanza-aprendizaje de la Inteligencia Artificial. Tesis en opción al grado de doctor en Ciencias de la Educación. Universidad Tecnológica de La Habana (CUJAE). Cuba
- Coca, Y., y Pérez, M. T. (2018).** Estudio sobre utilización de software libre educativo en el proceso de enseñanza-aprendizaje de la

Inteligencia Artificial del Ingeniero en Ciencias Informáticas. En: Ciencia e innovación tecnológica. Vol. II, Capítulo Ciencias Pedagógicas. LasTunas: EDACUN-Opuntia Brava.

**Coca, Y., & Pérez, M. T. (2020).** Integración de software libre educativo al proceso de enseñanza-aprendizaje en carreras de informática. *Referencia Pedagógica*, 8(1), 167–181.

*“Desarrollo y retos de la Inteligencia Artificial”* es el primero de cuatro libros que tienen como objetivo principal propiciar la incorporación de conceptos e ideas generales sobre Inteligencia Artificial en las escuelas. Se exponen algunos de los fundamentos de la Inteligencia Artificial, hechos relevantes que han marcado su desarrollo y su impacto en la sociedad. Se presenta un breve apartado sobre su enseñanza en Cuba y se brindan algunos elementos de su futuro y retos en el ámbito educativo. Se abordan los principales riesgos que amenazan a la humanidad si no se guardan principios éticos en el desarrollo de los sistemas inteligentes. Se presentan criterios y acciones específicas para lograr un desarrollo ético y sostenible de la Inteligencia Artificial.

Pretendemos que esta obra sirva de base para contribuciones teóricas y prácticas mediante el desarrollo de actividades educativas curriculares y no curriculares para las instituciones educativas, escuelas, universidades, familias, comunidades y la sociedad en general, así como para las instituciones y organizaciones gubernamentales y no gubernamentales; y de todos los sectores e instituciones de la sociedad. No se trata de enseñar a desarrollar aplicaciones de Inteligencia Artificial, sino, de contribuir al desarrollo del pensamiento computacional y nociones básicas necesarias para incentivar el estudio y profundización en IA con un posicionamiento ético, crítico y creativo, desde edades tempranas.

ISBN: 978-959-18-1340-4



9 789591 813404