

Inteligencia artificial y modelos conversacionales

Guía práctica para educación e
investigación

UAH OTRAS PUBLICACIONES

Inteligencia artificial y modelos conversacionales

Guía práctica para educación e
investigación

Jaime Oyarzo Espinosa
Luis Usero Aragonés
Francisco Javier Bueno Guillén



Universidad
de Alcalá

EDITORIAL
UNIVERSIDAD DE ALCALÁ

Licencia de esta obra:



Creative Commons Atribución/Reconocimiento-NoComercial-SinDerivados 4.0 Internacional <https://bit.ly/3RAKij4>

Puedes Compartir, copiar y redistribuir el material en cualquier medio o formato, bajo los siguientes términos:



Atribución: dar crédito de manera adecuada, brindar un enlace a la licencia, e indicar si se han realizado cambios.



NoComercial: no hacer uso del material con propósitos comerciales



SinDerivadas: Si remezclas, transformas o creas a partir del material, no podrás distribuir el material modificado.

Imágenes

Las Imágenes generadas por IA en esta obra están licenciadas como CC0 1.0 (a excepción de las explícitamente descritas que incluyen derechos de autor).

Normas APA (7ª edición) (Figuras – Normas APA, 2020), «Cuando la figura es de elaboración propia no es necesario agregar ningún tipo de declaración de derechos de autor. En APA se asume que es de autoría del propio autor».

Creative Commons (Timid Robot, 2023): «Si creas obras utilizando IA generativa, puedes seguir aplicando licencias CC a las obras creadas con el uso de esas herramientas y compartir tu trabajo de la forma que desees. La licencia CC que elijas se aplicará al trabajo creativo que aportes al producto final, aunque la parte producida por el propio sistema de IA generativa pueda no estar protegida por derechos de autor. Alentamos el uso de CC0 (Timid Robot, 2023) para aquellos trabajos que no impliquen un grado significativo de creatividad humana.»

Imágenes generadas y editadas por apps gratuitas de IA: Adobe Firefly, Canva, Pixelcut.

Este libro dispone de los siguientes sitios web institucionales:

[URL] <https://edunet.uah.es/wp/ia/>

Acceso a actualizaciones periódicas y materiales complementarios.

Formulario de registro para la suscripción a notificaciones sobre novedades y nuevas ediciones.

Recursos bibliográficos y documentales de apoyo

[URL] https://edunet.uah.es/wp/ia/?page_id=113

Blog académico complementario, para fomentar la reflexión crítica, la colaboración interdisciplinaria y el diálogo académico entre investigadores, docentes y estudiantes.

© De los textos: sus autores.

© De las imágenes: sus autores.

© Diseño de portada y contraportada: Jaime Oyarzo Espinosa

Plaza de San Diego, s/n

28801 Alcalá de Henares

www.uah.es

I.S.B.N. electrónico: 978-84-10432-83-3

Composición: Solana e Hijos, A. G., S.A.U.

Impreso en España

Presentación

¿Quiénes somos?

Somos un equipo de investigación dedicado al análisis y difusión de las potencialidades transformadoras de la inteligencia artificial IA en el ámbito educativo. El carácter disruptivo de alto impacto de la IA se erige como un catalizador para la innovación pedagógica y metodológica, capaz de redefinir los paradigmas tradicionales de enseñanza-aprendizaje.

Esta guía práctica tiene como objetivo dual: (1) sistematizar conceptos fundamentales para la comprensión crítica de los sistemas de IA generativa, y (2) proponer estrategias aplicadas que permitan solventar desafíos concretos en contextos académicos e investigativos. Su diseño pedagógico escalable está orientado tanto a docentes y estudiantes en etapas iniciales de alfabetización digital como a profesionales que buscan optimizar sus competencias en la implementación avanzada de estas tecnologías.

Mediante un enfoque teórico-práctico, esta guía busca constituirse en un recurso para el desarrollo de habilidades especializadas, la consolidación de marcos conceptuales robustos y la aplicación efectiva de inteligencia artificial generativa en la práctica docente, la producción académica y la investigación.



Objetivo de la guía práctica

Esta guía tiene como objetivo servir como una guía integral para la comprensión e implementación de la Inteligencia Artificial Generativa y los Modelos de Lenguaje (LLM) en contextos de educación superior e investigación académica. A través de un enfoque multidisciplinar práctico, buscamos familiarizar a docentes, investigadores y estudiantes con los conceptos fundamentales de estas tecnologías emergentes, proporcionando estrategias concretas para optimizar procesos de enseñanza-aprendizaje, personalizar experiencias educativas y potenciar la investigación académica.

Nuestro objetivo es ofrecer no solo fundamentos teóricos, sino también aplicaciones prácticas mediante ejemplos adaptables, plantillas funcionales y técnicas efectivas de ingeniería de prompt. El texto aborda de manera equilibrada las ventajas transformadoras de IA junto con sus limitaciones, desafíos éticos y consideraciones sobre privacidad y seguridad.

Aspiramos a que esta guía práctica se convierta en un recurso didáctico valioso que empodere a la comunidad académica para aprovechar responsablemente el potencial de IA generativa, fomentando la innovación pedagógica, el pensamiento crítico y el desarrollo de nuevas metodologías que respondan a los retos educativos contemporáneos.

Audiencia

Este guía está dirigido a docentes, investigadores y estudiantes de educación superior interesados en aplicar la inteligencia artificial generativa en contextos académicos y de investigación. Resulta especialmente valioso para profesionales de ciencias, educación y disciplinas afines que buscan integrar las tecnologías de IA en sus prácticas de enseñanza, investigación, desarrollo de materiales didácticos y gestión de proyectos académicos.

Terminología técnica

Los autores deciden mantener la terminología técnica en inglés y puntualmente complementarla con una traducción para facilitar la comprensión del contenido.

Utilizamos “inteligencia artificial” e “IA” para abreviar un amplio campo de tecnologías y aplicaciones prácticas, que incluye machine learning ML

(aprendizaje automático), Neural Networks NN (Redes Neuronales), Deep Neural Network DNN (NN con múltiples capas ocultas), Deep Learning DL (Aprendizaje profundo) y Large Language Models LLM (modelos lingüísticos).

Utilizamos la abreviatura “IA” por cuestiones prácticas, aunque reconocemos que IA no es realmente “artificial” (es creada y utilizada por humanos), ni “inteligente” (al menos en el sentido de la inteligencia humana).

El glosario, en Apéndices, aporta una definición en español a cada uno de estos términos técnicos. Reconocemos que existen otras versiones diferentes de cada traducción.

Resumen Ejecutivo

Esta guía práctica explora el potencial transformador de los Modelos Lingüísticos (LLM) y la Inteligencia Artificial Generativa (IAGen) en la educación, abarcando desde sus fundamentos y aplicaciones hasta consideraciones éticas, limitaciones, riesgos y, de manera central, la importancia del “prompt engineering” para optimizar su uso. Se destaca la necesidad de un enfoque centrado en el ser humano, la implementación de regulaciones y la validación pedagógica de estas tecnologías en el contexto educativo.



Resumen de capítulos

Presentación	• ¿Quiénes somos? • Objetivo de la guía práctica • Audiencia • Terminología técnica • Resumen Ejecutivo • Resumen de capítulos
Capítulo I Introducción	• Algunas observaciones previas • Herramientas de Inteligencia Artificial en la guía • Conceptos de IA generativa y Modelos Lingüísticos LLM • IA Generativa y conceptos relacionados • Evolución histórica desde Turing hasta el deep learning • El aprendizaje humano y el “aprendizaje” de IA • Comparación: Aprendizaje humano vs “aprendizaje” IA
Capítulo II Ética, Limitaciones y Riesgos	• Aspectos éticos relacionados con el uso de IA • Sesgos algorítmicos, privacidad y desinformación • Limitaciones y riesgos de los sistemas de IA generativa • Estrategias para mitigar riesgos y promover un uso responsable de IA • Marcos regulatorios (Ley Europea de IA) y casos de daños (deepfakes, discriminación)
Capítulo III Aplicaciones de Referencia	• Casos en medicina (diagnóstico predictivo, ahorro de costos), medio ambiente (gestión de incendios, optimización energética) • Aspectos positivos, beneficios, desafíos y consideraciones éticas y perspectivas futuras • Equilibrio entre innovación y sostenibilidad, con ejemplos en medicina y herramientas de reciclaje con IA.
Capítulo IV Fundamentos	• Tecnologías subyacentes de IA: Machine Learning redes neuronales, Deep Learning, Generative AI, modelos lingüísticos, procesamiento del lenguaje, chatbots y agentes IA. • AI en Educación e Investigación: IA Generativa está transformando la forma como enseñar, aprender, investigar y llevan a cabo la gestión académica. Consideraciones éticas centrales • Funcionamiento de los modelos lingüísticos. • Tokens y Longitud de contexto vinculada a las arquitecturas Transformer. • Tokenización, ventanas de contexto y limitaciones técnicas. • Tablas comparativas de modelos y longitud de contexto vs métodos para gestionar documentos extensos. • Capacidades de razonamiento de los LLM

Capítulo V - Prompt, Preguntas, Instrucciones	- Definición de prompt, como la instrucción textual que permite comunicar nuestras intenciones a una herramienta IA. - Diseño de prompt efectivos y aspectos clave a considerar: Contextualización, Especificidad de la respuesta, Gestión de la ambigüedad, Audiencia. - Ejemplo de estructura de prompt, rol asignado a IA, definición de la tarea, Requisitos, Instrucciones y Formato. - Diversos tipos de prompt, propósito y ejemplos descriptivos - Prompt efectivos para la docencia
Capítulo VI Ingeniería de prompt Prompt Engineering	- Ingeniería de prompt, una técnica para optimizar el rendimiento de los modelos de IA generativa. - Estrategias de ingeniería de prompt y ejemplos - Flujo de trabajo para una efectiva ingeniería de prompt - SYSTEM vs USER prompt - Marcos de prompt (prompt Frameworks), metodologías que garantizan prompt relevantes, eficientes y significativos - Ejemplos de prompt Frameworks: CARE, RICE, RISEN, FOCUS, SOCRATIC, BLOOM'S TAXONOMY, Chain of Thought CoT, Chain of Destiny, Playoff Method. - Símbolos y Etiquetas de prompt: caracteres especiales o palabras clave que proporcionan instrucciones adicionales a IA
Capítulo VII IA en Educación	- Aplicaciones académicas de Modelos Lingüísticos (LLM) y algunas estrategias educativas con IA - Métodos para integrar IA en evaluación (rúbricas dinámicas), creación de contenidos (resúmenes, infografías), ejercicios de combinación de IA y el desarrollo de pensamiento crítico. Analiza riesgos (plagio) y soluciones (actividades colaborativas, énfasis en habilidades humanas). ¿Cómo me ayuda IA en mi práctica docente y profesional?
Capítulo VIII IA en Investigación	- Descripción de herramientas de inteligencia artificial diseñadas para optimizar la investigación académica y la docencia. - Aspectos destacados de la guía publicada por UNESCO sobre el uso de la IA Generativa (IAGen) en educación e investigación. - Presentación de herramientas como Research Rabbit (Exploración de Literatura Académica), Elicit AI (Análisis y Comprensión de Artículos Científicos), Connected Papers (Visualización de Redes de Conocimiento), Scite (Evaluación del Impacto de Citas Científicas), Perplexity AI (Búsqueda Inteligente con IA), Jenni AI, (Asistencia en Escritura Académica) y traducción especializada con DeepL - Tabla comparativa de estas herramientas

Capítulo IX Apéndices	• Apéndice 1: Ventajas y Riesgos Potenciales de IA • Apéndice 2: Biblioteca de prompt • Apéndice 3: Otros Recursos (Artículos, E-Books) • Apéndice 4: Ofertas de autoaprendizaje gratuitos • Apéndice 5: Glosario de Términos IA • Apéndice 6: FAQ – Preguntas Frecuentes • Apéndice 7: Preguntas y Respuestas de esta Obra • Apéndice 8: Referencias Bibliográficas • Índice de Figuras • Índice de Tablas
--------------------------	---

CAPÍTULO I - INTRODUCCIÓN

Introducción

- Conceptos y fundamentos
- Ejemplos, «plantilla» para ser adaptados en cada caso
- Información de referencia, documentos y sitios web
- Recomendaciones para el aprendizaje complementario: «¿Deseas ampliar tu comprensión del tema?»
- Ventajas, riesgos y limitaciones inherentes a Large Language Models LLM (Modelos Lingüísticos conversacionales)
- Consideraciones éticas, privacidad y seguridad

Los LLM se encuentran en una fase de desarrollo y muestran un gran potencial transformador en una amplia variedad de aplicaciones.

Por ejemplo, el desarrollo de nuevas herramientas y recursos educativos, la mejora del servicio al cliente, la creación de nuevas formas de arte y entretenimiento, así como la automatización de tareas repetitivas en diversos sectores.

Los modelos lingüísticos (LLM), y la rápida evolución, plantea diversos retos y cuestiones éticas.

Una de las principales preocupaciones es el sesgo en los datos de entrenamiento, que puede dar lugar a respuestas injustas y estereotipos.

Otro riesgo es el uso de IA para amplificar ideologías nocivas o difundir información errónea.

Es necesario también considerar el impacto de los LLM en el empleo y la necesidad de garantizar un acceso equitativo a los programas de educación y formación.

Puedes encontrar la definición de los términos relacionado con IA mencionados en el texto en el apéndice 4: Glosario de términos IA.

¿Deseas ampliar tu comprensión del tema?

[¿Qué es Google Gemini? Guía](#)

[Bienvenido a Claude](#)

[ChatGPT e IA en la educación superior: guía de inicio rápido - Unesco](#)

¿Cómo comenzar?

Esta guía práctica es un recurso didáctico que describe las características y potencialidades de Large Language Models LLM (Modelos Lingüísticos conversacionales), para facilitar la comprensión y aplicación en ámbitos educativos y de investigación.

Algunas observaciones previas

- Las preguntas, propuestas de trabajo, de análisis y de investigación estarán orientadas a cualquier versión gratuita de LLM,
- Utilizaremos, para los ejemplos los siguientes LLM:
 - **Anthropic AI Claude** <https://claude.ai/>
 - **Google Gemini** <https://gemini.google.com>
 - **High-Flyer DeepSeek** <https://www.deepseek.com/>
 - **Microsoft Copilot** <https://copilot.microsoft.com/>
 - **OpenAI ChatGPT** <https://chatgpt.com/>
 - **Perplexity AI** <https://www.perplexity.ai/>

(Algunos LLM utilizados en varias consultas, no proporcionan una URL pública para compartir o publicar los resultados de los prompt.)

Preguntas:

- ¿Es la inteligencia artificial real apoyo a la práctica docente?
- ¿Cómo exploro sus ventajas y recomendaciones?
- ¿Cómo analizo las limitaciones, riesgos y desafíos?
- ¿Cómo me preparo?

¿Qué opina la propia inteligencia artificial?

Preguntemos a la Inteligencia artificial, acerca de sus recomendaciones, de sus ventajas, limitaciones, riesgos y desafíos en la aplicación educativa de la inteligencia artificial.

Incluyamos algunas restricciones a la pregunta: deseamos una respuesta breve, que especifique los primeros pasos para comenzar a utilizar IA, con una longitud de 1 párrafo y un tono académico y amistoso.

Exploremos las respuestas de: **ChatGPT, Claude y Gemini.**

Preguntemos: «Describe tus recomendaciones, ventajas, limitaciones, riesgos y desafíos en la aplicación educativa de la inteligencia artificial generativa. ¿Cómo comienzo a utilizar «LLM»? Responde de forma resumida, longitud de 1 párrafo, tono académico y amistoso». (la variable «LLM» es reemplazada por «ChatGPT», «Claude», «Gemini»)

Respuesta de Anthropic AI Claude

La implementación de inteligencia artificial generativa en educación ofrece ventajas significativas, como la mejora en la generación de ideas y la gestión de datos, pero también presenta desafíos importantes, incluyendo preocupaciones sobre la integridad académica y problemas de fiabilidad. Para comenzar a utilizar LLM de manera efectiva, se recomienda un enfoque estructurado que considere sus fortalezas y limitaciones únicas. Es crucial reconocer que, aunque la línea entre la innovación y el plagio puede ser difusa, una comprensión equilibrada de los riesgos y beneficios permitirá aprovechar esta tecnología como una herramienta complementaria en el entorno académico.

Respuesta de Google Gemini

La inteligencia artificial generativa en educación tiene un gran potencial para personalizar el aprendizaje y crear experiencias más interactivas. ¡Pero no es magia! Es importante ser consciente de sus limitaciones. Por ejemplo, IA no puede reemplazar la interacción humana en el aula. Para empezar con Gemini, ¡es muy fácil! Solo abre una conversación y comienza a preguntarme lo que quieras. A medida que interactúes conmigo, iré aprendiendo de tus necesidades y mejorando mis respuestas.

Respuesta de High-Flyer DeepSeek

La inteligencia artificial generativa (IA-G) puede personalizar el aprendizaje, generar contenido didáctico y ofrecer tutorías automatizadas, mejorando la eficiencia educativa. Su uso en corrección de textos, diseño de currículos y simulaciones interactivas es prometedor, como muestran herramientas como ChatGPT o Gemini. Limitaciones y Riesgos: Dependencia en datos de calidad, posibles sesgos algorítmicos y desafíos éticos (plagio, privacidad). Desafíos: Integrar IA-G sin reemplazar el rol docente, asegurar accesibilidad y validar su precisión. Para usar DeepSeek, regístrate en su plataforma, explora sus funciones (ej: resúmenes, Q&A) y aplícalo en actividades complementarias, siempre con supervisión crítica..

Respuesta de Microsoft Copilot

Para aplicar la inteligencia artificial generativa en educación, recomiendo una integración gradual y capacitación adecuada. Las ventajas incluyen personalización del aprendizaje y acceso continuo a recursos. Sin embargo, existen limitaciones como la dependencia tecnológica y los costos de implementación. Los riesgos abarcan la privacidad de datos y posibles sesgos en los algoritmos. Los desafíos incluyen la aceptación por parte de los usuarios y la necesidad de regulaciones claras. Para comenzar a utilizar Copilot, abre la aplicación, configura tu cuenta y explora sus funcionalidades. ¡Buena suerte en tu aventura educativa!

Respuesta de OpenAI ChatGPT

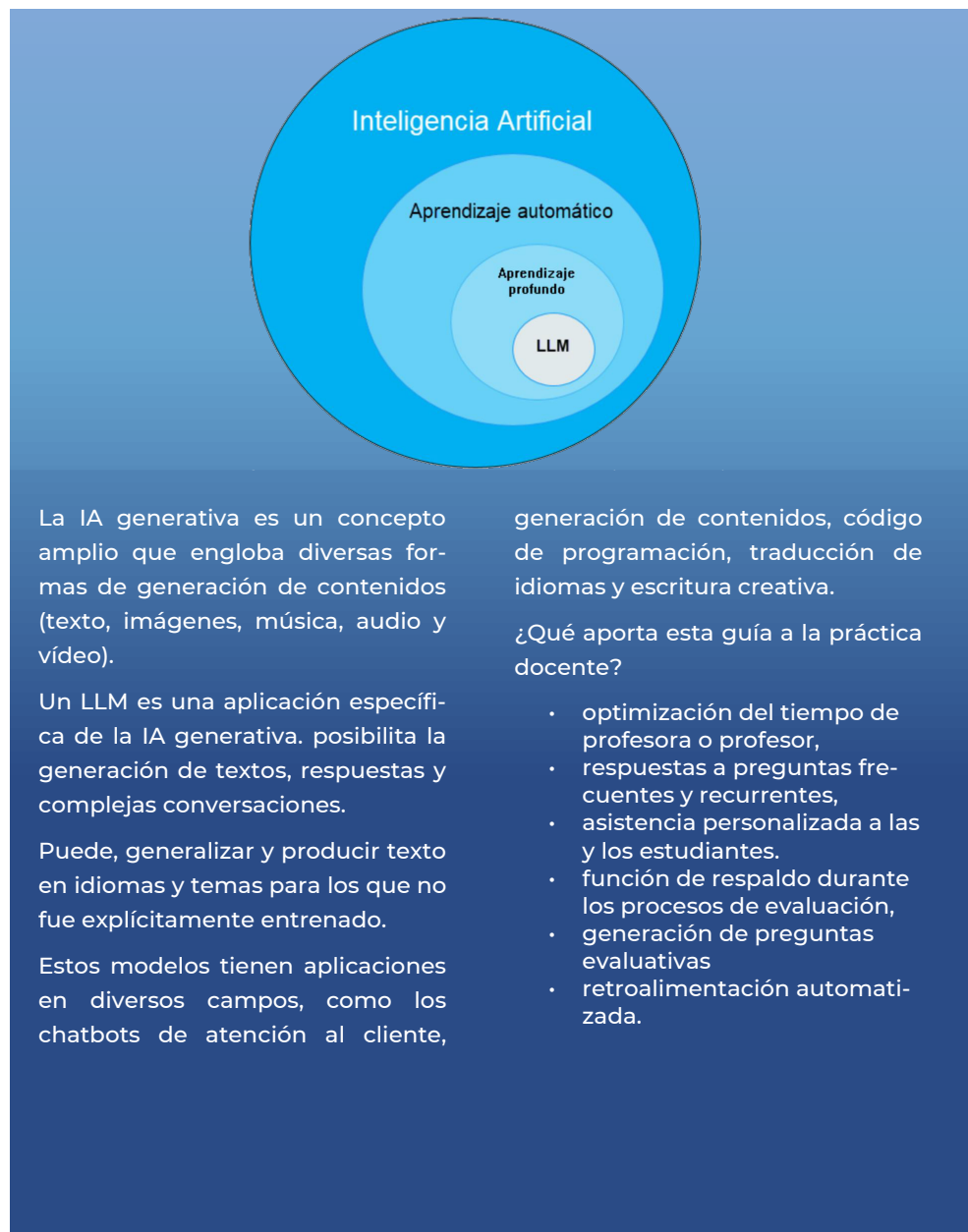
La aplicación educativa de IA generativa, como ChatGPT, ofrece ventajas significativas como personalización del aprendizaje, retroalimentación inmediata y acceso a vastos recursos de información. Sin embargo, enfrenta limitaciones como la falta de comprensión profunda del contexto, respuestas a veces inexactas y la posibilidad de generar contenido inapropiado. Los riesgos incluyen el uso excesivo de la tecnología, que puede limitar el pensamiento crítico, y la dependencia de datos sesgados. Comenzar a utilizar ChatGPT en la educación implica definir objetivos claros, explorar casos de uso pedagógico, y complementar su uso con la orientación de docentes para asegurar una experiencia educativa equilibrada y crítica

Respuesta de Perplexity

La inteligencia artificial generativa en educación ofrece ventajas como el aprendizaje personalizado, mayor motivación, eficiencia administrativa y acceso inclusivo, pero enfrenta limitaciones como la dependencia tecnológica, desinformación, pérdida de interacción humana y retos éticos relacionados con privacidad y equidad. Para aplicarla, recomiendo establecer directrices claras, fomentar el pensamiento crítico y supervisar su uso. Para comenzar con Perplexity, simplemente accede a la plataforma, explora sus funciones y utiliza preguntas precisas para obtener respuestas educativas relevantes y confiables.

Inteligencia artificial generativa

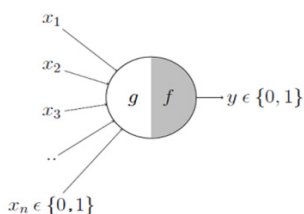
Figura 1. Elaboración propia, adaptación de (KAUR, 2024)



Breve historia de IA

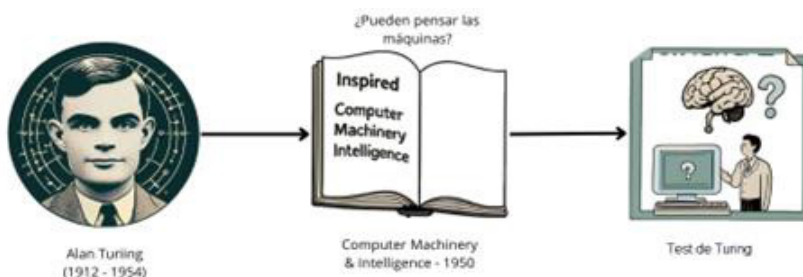
La Inteligencia Artificial IA comenzó a perfilarse a mediados del siglo XX. El primer gran hito remarcable fue un artículo escrito por McCulloch-Pitts en 1943 (McCulloch & Pitts, 1943), en el que, en la búsqueda de descifrar el funcionamiento del cerebro, diseña una unidad mínima de computación cerebral, que «paradójicamente» es la base de las redes neuronales artificiales. El siguiente diagrama es la expresión matemática de una neurona humana. Las señales se reciben a través de las dendritas (izquierda) y se envían a través del axón (derecha).

Figura 2. McCulloch-Pitts neuron (Interaction Design Foundation) (Neural Networks 2023)



Podemos datar el nacimiento de la Inteligencia Artificial en 1950, cuando Alan Turing, una de las mentes más brillantes, fue el primero en plantear que elemento mecánico podría tener habilidades humanas, si las máquinas podían pensar en algún momento, y en cómo demostrar ese hecho.

Figura 3. Alan Turing, IA, Test de Turing



La denominación «inteligencia artificial» fue acuñado en el verano de 1956. Un grupo de matemáticos decidió reunirse en la Conferencia de Dartmouth College in Hanover, New Hampshire, una pequeña localidad al norte de Boston, Marvin Minsky, Herbert Simon y Allen Newell y McCarthy

definen IA como «la ciencia y la ingeniería de crear máquinas inteligentes» y programas de computación inteligentes, relacionada con la utilización de ordenadores para comprender la inteligencia humana.

Figura 4. J. McCarthy, M. Minsky, N.I Rochester, C. Shannon, R. Solomonoff



Aunque Alan Turing no participó en la Conferencia de Dartmouth, los trabajos realizados por Turing (fallecido en 1954) desempeñaron un papel fundamental en el desarrollo de la Inteligencia Artificial.

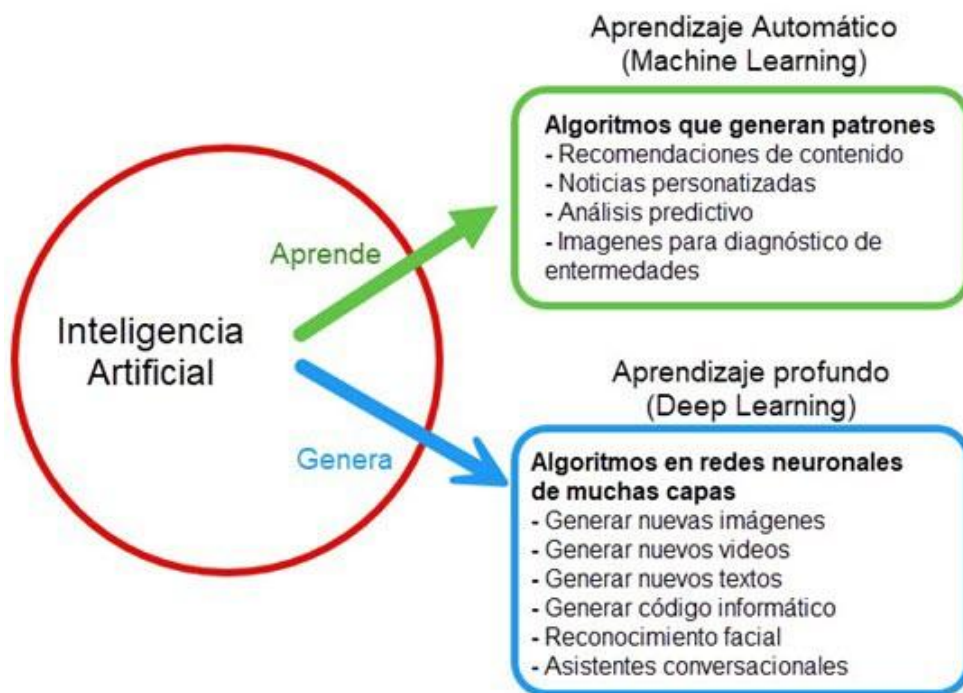
IA ha pasado por diferentes periodos de sensacionalismo al ostracismo, por diferentes razones, como se observa en la siguiente tabla.

Tabla 1. Períodos de IA

Período	Descripción
1943 – 1955	Primeros pasos excitantes
1956 - 1973	Comienzo extremadamente optimista
1974 - 1980	Primer Invierno
1981 - 1987	Recuperación
1988 - 1993	Primer Invierno
1994 - 2011	Relanzamiento
2012 - ???	Rentabilidad económica

Esta investigación tiene un repunte en los años 80 con la aparición de los sistemas expertos, programas informáticos capaces de actuar como un experto en algunos dominios específicos.

Figura 5. IA, Algoritmos y NN



¿LLM vs IA generativa?

Large Language Model LLM (Modelos Lingüísticos) y la Generative AI (IA Generativa) son conceptos relacionados de la inteligencia artificial. Pero son conceptos diferentes.

- **Large Language Model, LLM** es un tipo de IA diseñado para procesar y generar un lenguaje similar al humano. Se entrena con grandes conjuntos de datos para puede responder a preguntas, instrucciones, resumir textos e incluso generar contenidos creativos.
- **Generative AI**, un término más amplio que incluye a los LLM, pero también abarca otros modelos de IA diseñados para crear contenidos originales, como imágenes, música y vídeos.

Aunque ambos tipos de IA pretenden imitar la creatividad humana, los LLM se centran específicamente en el procesamiento y la generación de Natural Language Processing, NLP (lenguaje natural).



IA generativa y, en particular los LLM, se encuentra en una fase de desarrollo acelerado. Las capacidades, riesgos, limitaciones y mejores prácticas están en permanente evolución.

Chatbot y LLM

Un Chatbot o LLM se refiere a tecnologías avanzadas de inteligencia artificial (IA) que pueden mantener conversaciones con humanos de forma natural con ayuda del Natural Language Processing, NLP.

- **Chatbot** de IA: programa informático diseñado para simular conversaciones de tipo humano con los usuarios, normalmente a través de interacciones de texto o voz. Es como un asistente que nunca duerme ni se toma un café. Los chatbots pueden gestionar un gran volumen de consultas simultáneamente, reduciendo los tiempos de espera.
- **LLM:** modelo de IA específico, entrenado para asistir y proporcionar información a los usuarios. Por ejemplo: **ChatGPT, Claude, Gemini, Copilot, DeepSeek, Perplexity**. son los modelos IA empleados en esta guía práctica.

Ambas tecnologías (Chatbot y LLM), aprovechan el aprendizaje automático y el procesamiento del lenguaje natural (Natural Language Processing, NLP), para comprender y responder a la interacción humana, revolucionando la forma en que interactuamos con las máquinas y accedemos a la información.

A menudo, los términos **Chatbot o LLM** se utilizan indistintamente.

IA y aprendizaje humano

Conceptos diferentes de aprendizaje

En idioma inglés, a menudo se asocia la IA con capacidades de «aprender de la experiencia». Sin embargo, aunque en esta guía práctica se utilice el término «aprendizaje», este debe entenderse en un sentido técnico y no como un proceso análogo al aprendizaje humano. En el contexto de IA, el «aprendizaje» se refiere a la capacidad de los sistemas para mejorar su desempeño o adaptarse a través de la optimización matemática y métodos estadísticos, sin implicar conciencia, comprensión o conocimiento experimental.

Esta distinción es crucial, ya que el «aprendizaje» en IA no involucra experiencias subjetivas ni procesos cognitivos similares a los humanos, sino la identificación de patrones en datos y la ejecución de algoritmos diseñados para tareas específicas. Comprender esta diferencia es fundamental para evitar interpretaciones inexactas sobre las capacidades y limitaciones de IA.

Describir las capacidades de IA sin mencionar «aprender de la experiencia»

1. Procesamiento de datos y análisis

- Procesamiento de grandes volúmenes de datos para identificar patrones, relaciones y tendencias que podrían pasar desapercibidos para los humanos.
- Ejemplo: Análisis predictivo de resultados futuros basándose en datos históricos, predicción de tendencias del mercado.

2. Automatización de tareas

- Automatiza tareas repetitivas, peligrosas o tediosas, liberando tiempo para actividades humanas más complejas.
- Ejemplo: Verificación de documentos o transcripción automática.

3. Reconocimiento y clasificación

- Reconocimiento de imágenes, texto o voz y clasificación de elementos según criterios predefinidos.
- Ejemplo: Reconocimiento facial o conversión de voz a texto.

4. Interacción con lenguaje natural

- Procesamiento y generación de lenguaje humano para responder preguntas, traducir textos o mantener conversaciones.
- Ejemplo: Asistentes virtuales como Alexa o Google Assistant.

5. Optimización basada en retroalimentación

- Ajuste del rendimiento mediante técnicas como aprendizaje supervisado o por refuerzo, optimizando decisiones a partir de recompensas o penalizaciones.

6. Velocidad y escalabilidad

- Procesa información y gestiona grandes volúmenes de datos más rápido que los humanos.

7. Disponibilidad

- Funciona 24/7 sin interrupciones.

8. Adaptabilidad

- Mejora su desempeño al ser expuesta a nuevos datos mediante ajustes en sus algoritmos.

Resumen capítulo I: Introducción

Una introducción a los Modelos de Lenguaje (LLM) y la Inteligencia Artificial Generativa en el contexto educativo. Comienza estableciendo los fundamentos conceptuales de estas tecnologías, distinguiendo entre LLM como sistemas especializados en procesamiento de lenguaje y la IA Generativa como un campo más amplio, denominado multimodal, que incluye la creación de diversos tipos de contenidos.

El texto ofrece una perspectiva histórica del desarrollo de la IA, destacando hitos como el trabajo de McCulloch-Pitts sobre neuronas artificiales (1943) y las contribuciones de Alan Turing. Menciona la Conferencia de Dartmouth (1956) como el momento en que se acuñó el término "inteligencia artificial" y describe los ciclos de optimismo y desencanto que ha experimentado el campo.

Se enfatiza la distinción entre el "aprendizaje" en sistemas de IA y el aprendizaje humano, aclarando que el primero se basa en optimización matemática y métodos estadísticos sin implicar conciencia o comprensión. El documento enumera las capacidades reales de los sistemas de IA: procesamiento de datos, automatización, reconocimiento de patrones, interacción mediante lenguaje natural, optimización y adaptabilidad.

CAPÍTULO II - ÉTICA, LIMITACIONES Y RIESGOS

ÉTICA, LIMITACIONES Y RIESGOS

Aspectos Éticos de IA

El rápido auge de la inteligencia artificial (IA) ha generado nuevas oportunidades a nivel global: desde facilitar los diagnósticos de salud hasta posibilitar las conexiones humanas a través de las redes sociales, así como aumentar la eficiencia laboral mediante la automatización de tareas.

Sin embargo, estos rápidos cambios también plantean profundos dilemas éticos, que surgen del potencial que tienen los sistemas basados en IA para reproducir prejuicios, contribuir a la degradación del clima y amenazar los derechos humanos, entre otros. Estos riesgos asociados a IA se suman a las desigualdades ya existentes, perjudicando aún más a grupos históricamente marginados.

UNESCO (UNESCO, n.d.)

Los algoritmos de IA, diseñados para procesar datos y prever patrones, no solo toman decisiones, sino que también moldean opiniones y preferencias, especialmente en el ámbito digital. Esto genera inquietudes sobre la información alineada con creencias preexistentes, que limita la diversidad de perspectivas y socava el pensamiento crítico. La capacidad de estos algoritmos para modificar comportamientos plantea cuestiones éticas sobre la responsabilidad de empresas y desarrolladores en la promoción de patrones adictivos, afectando particularmente a grupos vulnerables como los menores.

Figura 6. Neurorights Foundation (The Neurorights Foundation, 2025)



- **Errores y fallos técnicos:** Los sistemas de IA pueden cometer errores debido a limitaciones en los datos de entrenamiento, algoritmos imperfectos o condiciones de uso no previstas. Estos errores o fallos afectan la confiabilidad en aplicaciones críticas.

Estos riesgos y limitaciones exigen marcos regulatorios robustos y un diseño responsable de la tecnología.



Potenciales Riesgos y Daños

Tabla 2. Riesgos y daños, basado en OECD (OECD Legal Instruments, 2025)

TIPO DE DAÑO	EJEMPLOS	COMO MITIGAR EL RIESGO	PREGUNTAS ABIERTAS
Daño intencionado	Vídeo deepfake diseñado para dañar la reputación y el derecho a la intimidad de una persona	Inversiones en tecnología de detección	¿Cómo evitar los deepfake perjudiciales sin coartar la libertad de expresión?
Daños causados por “efectos secundarios”	Datos de entrenamiento sesgados que conducen a la discriminación Algoritmo de redes sociales que fomenta el discurso del odio o la información falsa	Revisión de riesgos durante el desarrollo del producto Mayor transparencia para el usuario	¿Qué controles son necesarios para reducir estos riesgos en los sistemas de IA? ¿Deberían las empresas de IA ofrecer mayor transparencia sobre los riesgos de sesgo o incitación al odio? ¿Papel de los gobiernos?
Daños causados por porcentajes de fallos	Falsos positivos por reconocimiento facial en las fuerzas de seguridad	Desarrollo de productos con debida atención en materia de derechos humanos del cliente	¿Cómo mitigar los riesgos de daños causados por fallos de IA? ¿Qué herramientas y sistemas se necesitan para garantizar una reparación adecuada y asegurar que se aprenda de los errores?

TIPO DE DAÑO	EJEMPLOS	COMO MITIGAR EL RIESGO	PREGUNTAS ABIERTAS
Daños causados por un uso mal intencionado	Desinformación política mediante datos/sistemas de IA	Atención y diligencia de clientes y usuarios	¿Deberían las empresas de IA ofrecer mayor transparencia sobre el mal uso de IA para difundir desinformación u otros perjuicios? ¿Papel de los gobiernos?
Daños causados por fallos de seguridad	Hackers que toman el control de vehículos autónomos	Políticas públicas el desarrollo de productos	¿Cuál es el nivel de atención adecuado para garantizar la seguridad?

Resumen capítulo II: ética, limitaciones y riesgos

- Existen "profundos dilemas éticos" derivados del uso de la IA, incluyendo el potencial de "reproducir prejuicios, contribuir a la degradación del clima y amenazar los derechos humanos".
- Ejemplo de riesgos específicos como los sesgos algorítmicos, la falta de transparencia, la violación de la privacidad y la difusión de desinformación.
- El impacto potencial de la IA en el empleo subraya la necesidad de formación para adaptarse a estos cambios.
- Importancia de "estrategias para mitigar riesgos y promover un uso responsable de IA". Se menciona el "Marco regulatorio y Ley de Inteligencia Artificial de la Unión Europea", propuesta en abril de 2021.
- Mención a las Alucinaciones, los Sesgos, Privacidad de los datos, ética y responsabilidad, errores y fallos técnicos como imitaciones y riesgos centrales de IA.