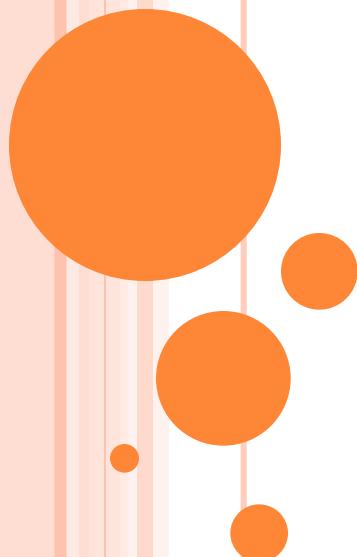


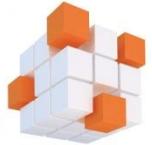
UNIVERSIDAD NACIONAL DE LA MATANZA

INTELIGENCIA DE NEGOCIOS

**Fundamentos de
Inteligencia Artificial**

Docente: ING. LORENA R. MATTEO





ROMPIENDO EL HIELO

The screenshot shows a presentation slide titled "Fundamentos de Inteligencia Artificial". On the left, there is a QR code with the text "Join at: ahaslides.com/SEG5P". The main content features a large blue "AI" text against a dark background with abstract blue and white geometric shapes. The AhaSlides logo is in the top right corner, and a navigation bar with icons is at the bottom.

**Fundamentos de
Inteligencia Artificial**

¿Qué sabes de IA?

INTELIGENCIA DE NEGOCIOS
EFC - UNLaM
Ing. Lorena R. MATTEO

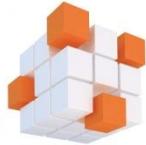
0/50



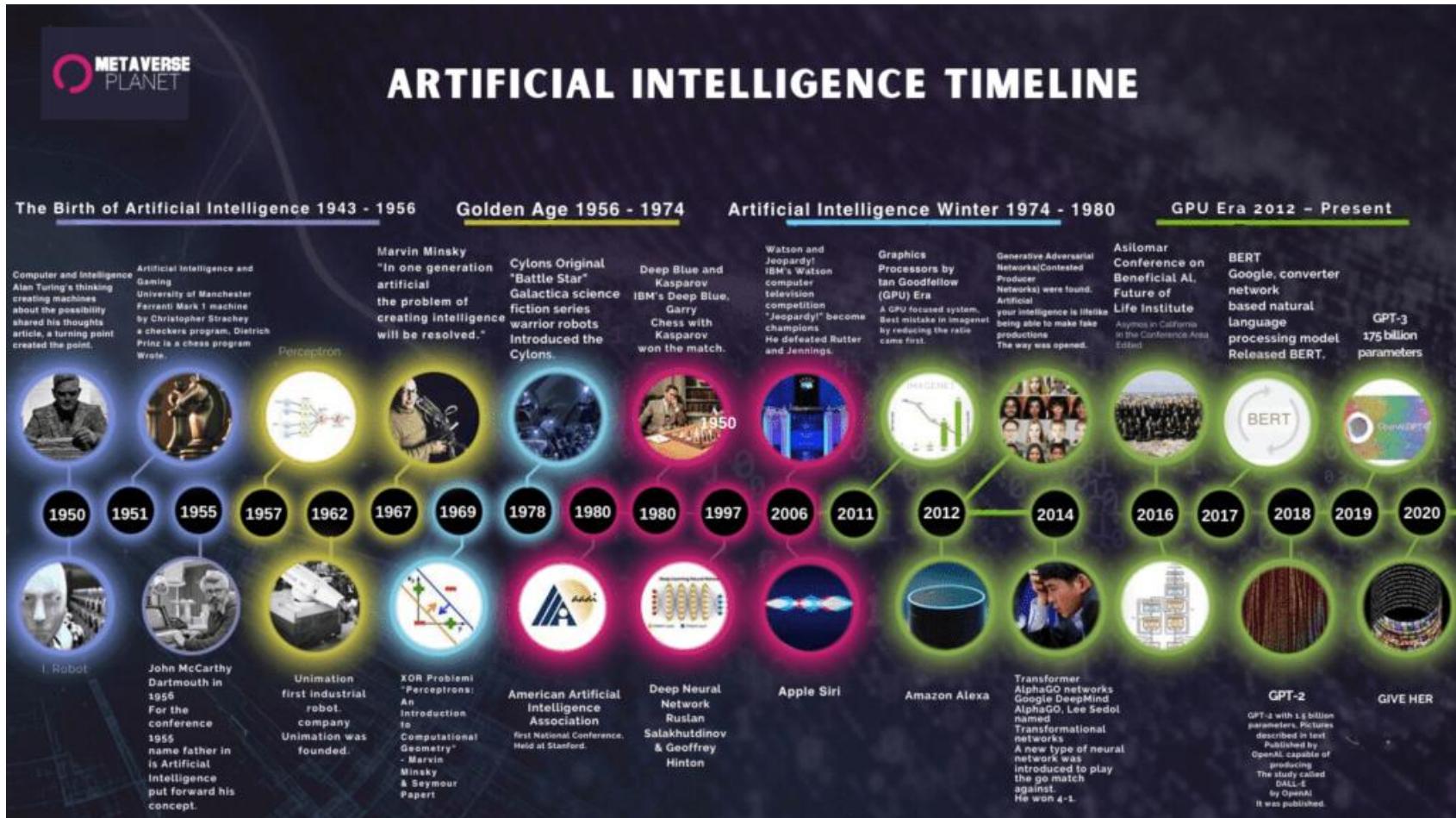
HISTORIA DE LA INTELIGENCIA ARTIFICIAL (RESUMEN CRONOLÓGICO)

- **Pre-1950s: Ideas Fundacionales y Lógica Simbólica:**
 - Desarrollo de la lógica formal por figuras como Aristóteles y Boole.
 - Invención de máquinas calculadoras y los primeros autómatas.
 - Trabajo de Alan Turing sobre la "Máquina de Turing" y el concepto de computabilidad.
- **Década de 1950: El Nacimiento de la IA:**
 - Taller de Dartmouth (1956) se considera el evento fundacional de la IA como campo de estudio.
 - Primeros programas de IA como el "Logic Theorist" y el "General Problem Solver".
- **Década de 1960: Optimismo y Primeros Éxitos:**
 - Desarrollo de programas que podían resolver problemas de álgebra, jugar al ajedrez y comprender el lenguaje natural en contextos limitados (e.g., ELIZA).
 - Predicciones optimistas sobre el rápido avance de la IA.
- **Década de 1970: El Primer "Invierno" de la IA:**
 - Desilusión por la falta de progreso significativo en la resolución de problemas complejos del mundo real.
 - Informe Lighthill critica la investigación en IA y lleva a la reducción de la financiación.
- **Década de 1980: El Renacimiento con los Sistemas Expertos:**
 - Auge de los sistemas expertos, diseñados para emular el conocimiento de expertos humanos en dominios específicos.
 - Aplicaciones en áreas como diagnóstico médico y configuración de computadoras.
- **Finales de los 80 y Década de 1990: El Segundo "Invierno" de la IA:**
 - Las limitaciones de los sistemas expertos y la falta de aplicaciones generalizadas llevan a una nueva disminución del interés y la inversión.
- **Finales de los 90 y Principios de los 2000: El Auge del Machine Learning:**
 - Mayor disponibilidad de datos y potencia computacional impulsan el desarrollo de algoritmos de aprendizaje automático.
 - Éxitos en áreas como el filtrado de spam, sistemas de recomendación y reconocimiento de voz.
- **Década de 2010: La Revolución del Deep Learning:**
 - Avances significativos en el aprendizaje profundo (Deep Learning) gracias a redes neuronales con múltiples capas.
 - Grandes avances en visión por computadora, procesamiento del lenguaje natural y reconocimiento de voz.
- **Presente (Década de 2020): La Era de los LLMs y la IA Generativa:**
 - Desarrollo y proliferación de modelos de lenguaje extensos (LLMs) como GPT-3 y modelos de IA generativa para imágenes, audio y video.
 - Debates sobre las implicaciones éticas, el potencial y los desafíos de la IA avanzada.

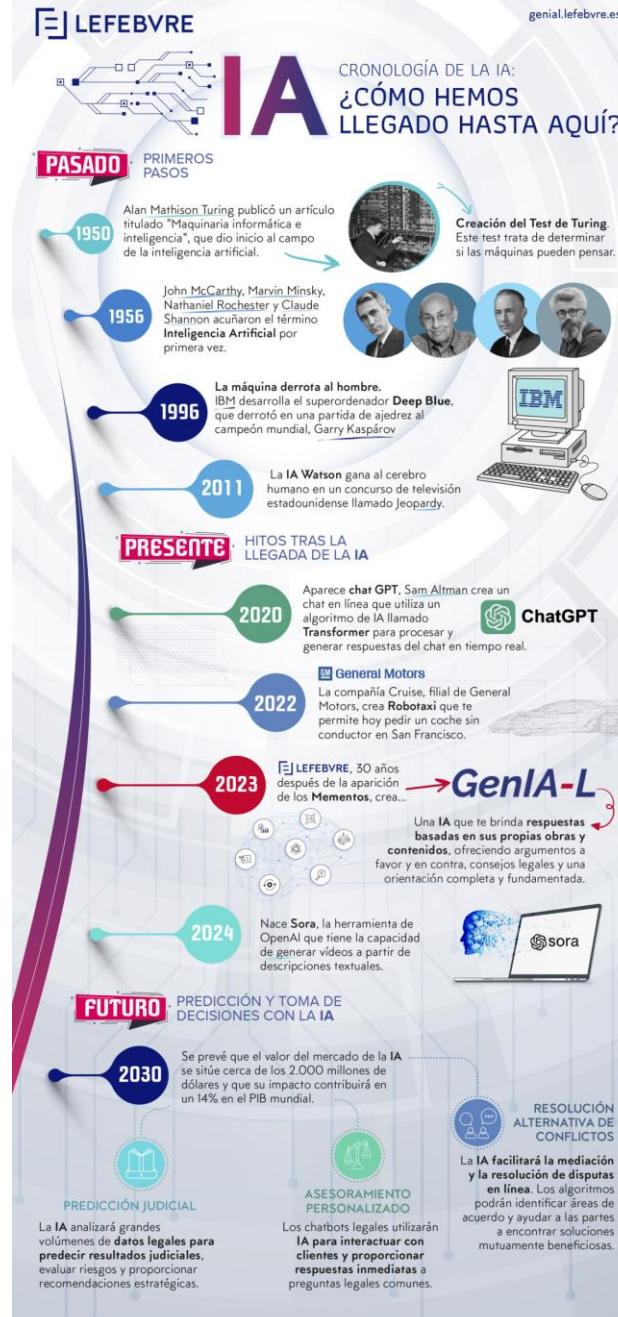
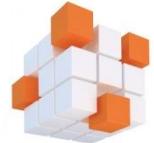
Fuente: Google. (2025). Gemini (versión del 10 de Junio) [LLM]. <https://gemini.google.com/>



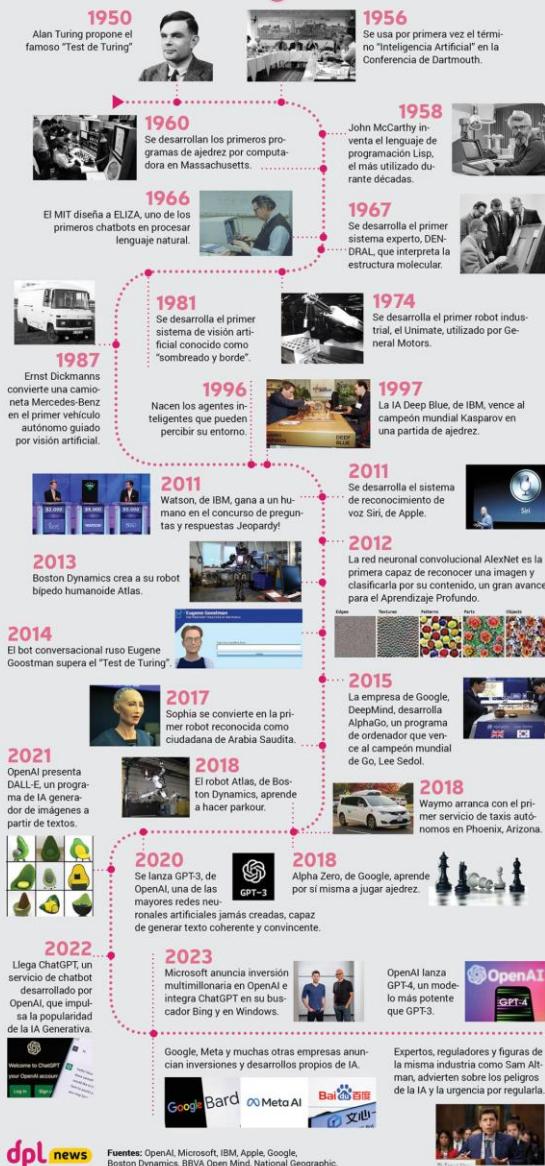
LÍNEA DE TIEMPO DE LA INTELIGENCIA ARTIFICIAL



Fuente: <https://metaverseplanet.net/blog/the-history-of-artificial-intelligence/>



La Inteligencia Artificial a través del tiempo



Fuente: <https://dplnews.com/infografia-la-inteligencia-artificial-a-traves-del-tiempo/>
<https://elderecho.com/cronologia-de-la-ia-como-hemos-llegado-hasta-aqui>



¿QUÉ ES LA INTELIGENCIA ARTIFICIAL GENERATIVA?

18/6/2025 IN2025

- La **Inteligencia Artificial (IA)** imita el comportamiento humano mediante el **aprendizaje automático** para interactuar con el entorno y ejecutar tareas sin instrucciones explícitas sobre qué generar.
- La **Inteligencia Artificial Generativa (IAGen)** describe una categoría de funcionalidades dentro de la inteligencia artificial que crean contenido original.



Diagram illustrating the process: Instruction → M-LLM → Response
- Estas funcionalidades incluyen **tomar entradas de lenguaje natural (Prompts)** y **devolver respuestas adecuadas en diversos formatos**, como **lenguaje natural, imágenes, código, etc.**

Fuente: [Introducción a los conceptos de inteligencia artificial generativa - Training | Microsoft Learn](#)
[Cómo funcionan los LLM y la inteligencia artificial generativa - .NET | Microsoft Learn](#)



¿QUÉ ES UN LLM?

- Un **LLM (Large Language Model)** es una forma avanzada de IA entrenada con grandes volúmenes de texto, que identifica patrones lingüísticos y genera lenguaje coherente.
- Los LLM descomponen el **texto en tokens**, palabras analizan el conocimiento adquirido y predicen el siguiente token para formar respuestas relevantes.
 - Dependiendo del modelo, un **token** puede representar una letra, una sílaba, una palabra completa o incluso un espacio en blanco.

A screenshot of a text editor interface. The main area displays a sentence: "Piensa como un asistente con 10 años de experiencia, organiza una reunión de 2 horas para revisar proyectos, discutir desafíos y planificar nuevas iniciativas del servicio universitario. Incluye al director, responsables y técnicos. Crea una agenda profesional y clara, con los temas, responsables y resultados esperados." The text is color-coded by token, with each word appearing in a different hue. Below the text, there is a table:

Tokens	Characters
79	321

Ejemplo de como un LLM descompone una frase en tokens para poder procesarla.



¿CÓMO FUNCIONA UN LLM?

“Las aplicaciones de IAGen que usamos hoy son posibles mediante la arquitectura transformer.”

“Los usuarios interactúan con la IAGen a través de aplicaciones de chat que usan lenguaje natural como entrada (Prompts).”

El Modelo:

- Descompone el texto en tokens (palabras, fragmentos o caracteres).
- Predice el siguiente token usando patrones y conocimientos previos.
- Construye respuestas coherentes y contextuales, como conectar puntos en un dibujo.

Claves:

- **Secuencia y coherencia:** Cada palabra depende de las anteriores.
- **Patrones ocultos:** Detecta patrones complejos en los datos.
- **Instrucciones claras:** Prompts bien estructurados guían respuestas precisas.
- **Influencia de la info previa en el LLM:** Las respuestas del LLM dependen de la coherencia y calidad de la información previa proporcionada.

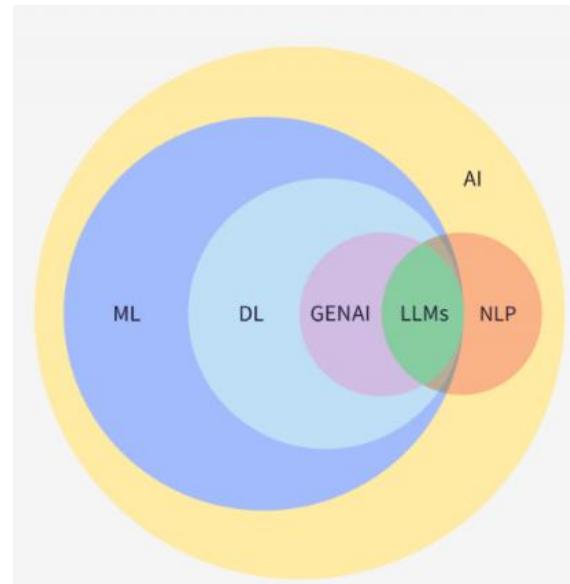


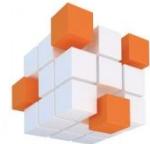
¿QUÉ SIGNIFICA DE GRAN ESCALA?

- Que el modelo tiene una cantidad enorme de **parámetros** (**neuronas artificiales**) y ha sido entrenado con enormes volúmenes de texto.
- Esta escala permite que pueda realizar tareas muy variadas, con respuestas cada vez más ajustadas y creativas.

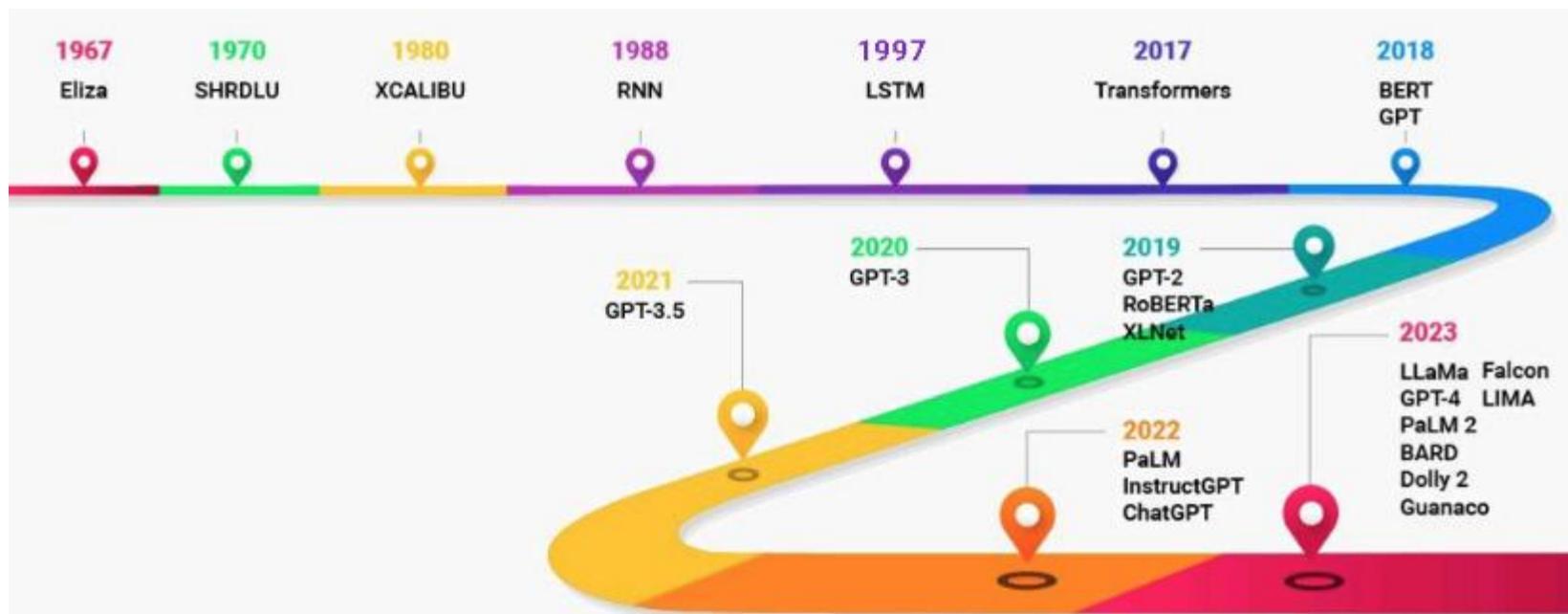
Ejemplos de LLMs conocidos:

- ChatGPT, DALLE3, SORA (OpenAI)
- Gemini, Veo2 (Google)
- Copilot (Microsoft)
- Claude (Anthropic)
- Mistral, LLaMA (Meta AI)
- Qwen, DeepSeek, entre otros.





EVOLUCIÓN LLMs





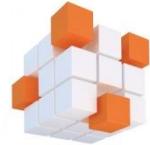
¿QUÉ ES UN SLM?

- Un SLM (**Small Language Model**) son versiones reducidas y especializadas de los LLMs.
- **Especialización:** están optimizados para tareas o dominios específicos.
- **Velocidad y personalización:** Se pueden ajustar y desplegar más rápido que los LLM.
- **Eficiencia computacional y menor huella energética:** Requieren menos recursos para entrenamiento e inferencia, lo que reduce costos y permite su ejecución en dispositivos con recursos limitados. Consumen menos energía, contribuyendo a una IA más sostenible.



COMPARACIÓN LLMS VS SLMs

Modelos de lenguaje grande (LLM)	Modelos de lenguaje pequeño (SLM)
Los LLM se entrena con grandes cantidades de texto que representan una amplia gama de temas generales, normalmente mediante el suministro de datos de Internet y otras publicaciones disponibles con carácter general.	Los SLA se entrena con conjuntos de datos más pequeños y centrados en el asunto
Cuando se entrena, tienen muchos miles de millones (incluso billones) de parámetros (pesos que se pueden aplicar a las incrustaciones vectoriales para calcular las secuencias de token previstas).	Normalmente, tienen menos parámetros que los LLMs.
Capaz de mostrar funcionalidades completas de generación de lenguajes en una amplia gama de contextos conversacionales.	Este vocabulario centrado hace que sean efectivos en temas conversacionales específicos, pero menos efectivos en la generación de lenguajes más general.
Su gran tamaño puede afectar a su rendimiento y dificultar la implementación local en dispositivos y equipos.	El tamaño menor de los SLM puede proporcionar más opciones para la implementación, incluida la implementación local en dispositivos y equipos locales; y hace que sean más rápidos y fáciles de ajustar.
Ajustar el modelo con más datos para personalizar su experiencia en el tema puede ser lento y costoso en términos de la potencia de proceso necesaria para realizar el entrenamiento adicional.	El ajuste fino puede ser potencialmente menos lento y costoso.



COMPARACIÓN LLMS vs SLMs

18/6/2025 IN2025

Quality vs Size in Small Language Models (SLMs)

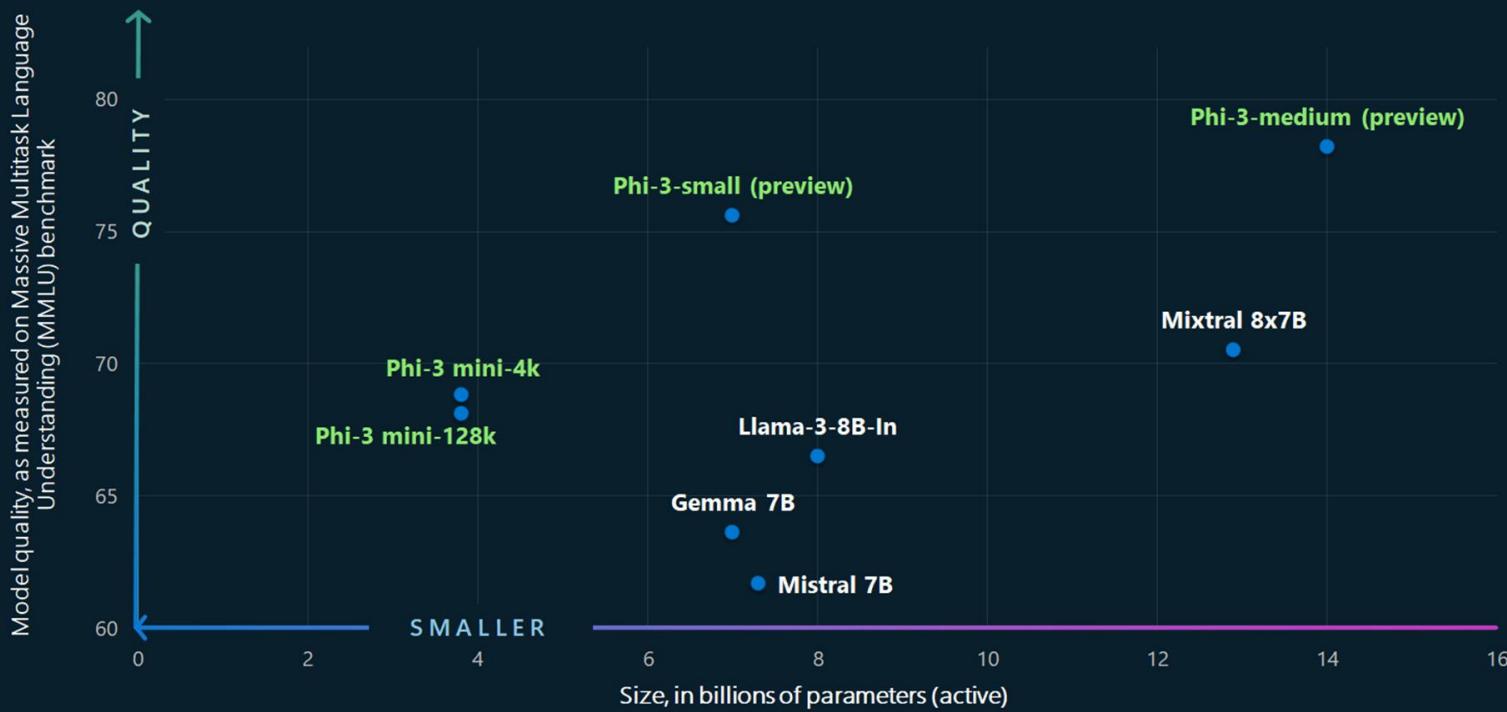
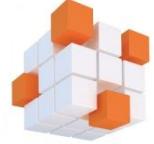


Gráfico que ilustra cómo la calidad de los nuevos modelos Phi-3, medida por el rendimiento en el benchmark Massive Multitask Language Understanding (MMLU), se compara con otros modelos de tamaño similar.
(Imagen cortesía de Microsoft)

Fuente: [Tiny but mighty: The Phi-3 small language models with big potential - Source](#)



¿QUÉ ES UNA ALUCINACIÓN?

- Aunque estos modelos de IAgen generan respuestas coherentes, no siempre son verídicas (**Alucinación**) y pueden reproducir **sesgos**, requiriendo revisión humana.

Cuando un modelo genera contenido falso pero con confianza y coherencia. También llamado **confabulación o delirio artificial**.

¿POR QUÉ OCURREN LAS ALUCINACIONES?

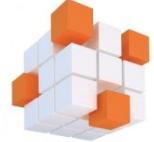
- Entrenamiento estadístico sin comprensión semántica o con errores de aprendizaje.
- Falta de verificación de los hechos.
- Prompts ambiguos o mal estructurados.



Principio clave: Usar la IA con pensamiento crítico y verificando fuentes.



DETECTAR EL RESULTADO ERRÓNEO



- Determinar cuál de los dos textos contiene cierta alucinación de la IA:

A

ELIZA fue un programa de procesamiento del lenguaje natural desarrollado en el MIT por Joseph Weizenbaum entre 1964 y 1966. Su objetivo era simular una conversación con un psicoterapeuta rogeriano. ELIZA funcionaba analizando las declaraciones del usuario en busca de palabras clave y luego generando respuestas basadas en esas palabras clave, a menudo reformulando las preguntas del usuario o utilizando frases preestablecidas. Aunque a menudo se la describe como un ejemplo temprano de inteligencia artificial, Weizenbaum mismo enfatizó que ELIZA no entendía realmente el lenguaje ni los problemas de los usuarios.

B

ELIZA fue un programa de inteligencia artificial creado por Marvin Minsky en la Universidad de Stanford en la década de 1970. Este innovador programa fue diseñado para imitar las respuestas de un consejero psicológico, utilizando técnicas de procesamiento del lenguaje natural para analizar las entradas del usuario y generar respuestas coherentes. ELIZA fue uno de los primeros intentos de crear una máquina capaz de mantener una conversación significativa con humanos, marcando un hito importante en el desarrollo de la IA conversacional.



RESOLUCIÓN ALUCINACIONES

- El **Texto B** contiene la alucinación:

Errores:

- Error 1: Autor y Universidad: ELIZA no fue creado por Marvin Minsky, sino por **Joseph Weizenbaum**. Tampoco fue desarrollado en la Universidad de Stanford, sino en el **MIT (Instituto Tecnológico de Massachusetts)**.
- Error 2: Fecha: ELIZA fue desarrollada entre **1964 y 1966**, no en la década de 1970.

Justificación:

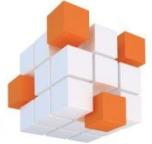
- *Aunque el Texto B describe correctamente la función general de ELIZA (simular un consejero psicológico utilizando PNL) y su importancia en la IA conversacional, atribuye incorrectamente su creación a Marvin Minsky y la ubica en la Universidad de Stanford en una década posterior. Estos detalles son factualmente incorrectos y constituyen una alucinación sutil, ya que podría parecer creíble para alguien que no conoce la historia específica de ELIZA.*



MEJORA DE LOS RESULTADOS DE LA SOLICITUD

- Un **Prompt** es una instrucción de texto clara y específica que guía la respuesta de un modelo de IA.
 - una pregunta, una consigna, una orden, una frase incompleta

“Es la llave que abre la caja de respuestas.”
- Para interactuar eficazmente con un LLM, se deben usar instrucciones claras, proporcionar contexto, dividir tareas complejas y dar ejemplos concretos.
- La **estructura de un buen prompt** incluye la tarea, el contexto, ejemplos, la persona, el formato y el tono deseado para la respuesta
 - Si es vago o ambiguo, el modelo se pierde o completa con suposiciones.
 - Si es claro, específico y detallado, da respuestas mucho más útiles.

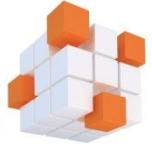


MIEL – Crear prompts efectivos para LLMs

MIEL = Meta + Identidad + Estilo + Lenguaje

- **Meta:** ¿Para qué lo quiero?
- **Identidad:** ¿Desde qué rol responde la IA?
- **Estilo:** ¿Con qué tono o forma?
- **Lenguaje:** ¿En qué formato entregará la respuesta?

MIEL es uno de los acrónimo que ayudará a construir prompts claros, situados y eficaces, especialmente útil en educación, comunicación, diseño o investigación.



INGENIERÍA DE PROMPTS

Ejemplo Prompt MIEL

- Prompt Básico:

“Contame qué es la memoria en psicología.”

- Prompt MIEL:

“Explicame qué es la memoria en psicología, como si fueras un docente universitario, en tono claro y reflexivo, para una presentación breve de 2 minutos.”



INGENIERÍA DE PROMPTS

Otro Tip para un Buen Prompt

- **Primero:** contexto claro.
- **Después:** la orden específica.

Ejemplo:

Usar formato Markdown

Contexto

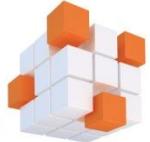
Estoy creando un informe para dirección
sobre la retención mensual de clientes.

Solicitud

Resume los 3 KPIs principales y su
evolución en formato tabla.

Material Adicional: [Markdown Guide](#)

Ver Guías en MIeL / Libros y Artículos Parte 2 - Minería de Datos y Explotación de Información:
"ChatGPT guía 2025 cómo crear Prompts como un profesional"

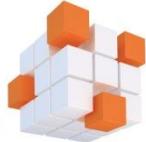


IA EN EL ÁMBITO EDUCATIVO

¿Para qué se usa en educación?

Aplicaciones útiles y Potencial educativo de la IA

- Resumir textos complejos.
- Generar ideas o actividades.
- Simular debates o corregir textos.
- Traducir o reformular contenidos.
- Crear consignas o materiales de clase.
- Asistencia en redacción.
- Corrección y sugerencias.
- Apoyo a la investigación.
- Evaluación automática.



IA EN EL ÁMBITO EDUCATIVO

Limitaciones y riesgos

- Pérdida de pensamiento crítico.
- Pérdida de autoría genuina del estudiante.
- Dificultad para evaluar autoría.
- Respuestas erróneas o inventadas (alucinación).
- Reproducción de sesgos o prejuicios.
- Falta de políticas institucionales claras

Claves para una integración crítica

- No es estar “a favor o en contra”.
- Promover uso reflexivo y con acompañamiento docente.
- Evaluar la comprensión, no solo el texto final.
- Acompañar con marcos éticos y pedagógicos.



CONSTRUIR MODELOS DE IAGEN

hadsonpar

¿Qué se necesita para construir modelos de IA Generativa?



1. Definir el Objetivo:

Entender y clarificar el propósito de del modelo generativo a construir.

IA Generativa



2. Selecciona un Framework:

Seleccionar un Framework de trabajo relacionado al aprendizaje profundo.



3. Recopilar datos de entrenamiento:

Recolectar un conjunto de datos relevante y diverso para entrenar el modelo.



4. Preprocesamiento de datos:

Aplicar el preprocesamiento necesario a los datos recopilados.



5. Selecciona la arquitectura del modelo Generativo:

Seleccionar la arquitectura generativa que más se acomoda al modelo.



6. Gestión de Recursos y Hardware:

Dimensionar la capacidad y la complejidad del modelo



7. Diseñar la arquitectura del modelo:

Diseñar la arquitectura del modelo, definiendo la estructura de las capas.



8. Desarrollo y entrenar del modelo:

Desarrollar el modelo seleccionado y entrenar en base a los datos seleccionados



9. Optimización del modelo:

Aplicar las técnicas de optimización para mejorar la eficiencia y el rendimiento del modelo,



10. Validación y evaluación:

Evaluar el rendimiento del modelo utilizando dataset de validación



11. Generación de nuevo contenido:

Continua actualización y mantenimiento del modelo Generativo.



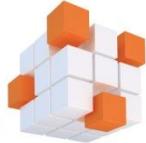
12. Consideraciones éticas y de privacidad:

asegurar las consideraciones e implicaciones éticas y de privacidad.



13. Ajustes y despliegue:

Realizar ajustes finales, como la optimización de inferencia y la gestión de recursos.



APLICACIONES – AUTOMATIZACIÓN CON IA

- Analizar una base de datos y generar insights clave**
→ ChatGPT Code Interpreter / Flowise AI / Hex.tech
- Explorar tendencias o temas de usuario para UX Research**
→ Perplexity AI / UserInterviews + Notion AI / ChatGPT
- Redactar y refinar contenido UX, copies o documentación**
→ ChatGPT + Grammarly + Jasper (free trial)
- Crear mapas de viaje o user flows a partir de texto**
→ Whimsical + [diagrams.net](#) + Mermaid.live editor
- Generar variantes de UI o wireframes básicos automáticamente**
→ Uizard.io / Galileo AI
- Organizar y priorizar tareas de producto o desarrollo**
→ Notion AI + Reclaim + Linear (con IA)
- Generar clips o tutoriales rápidos para productos**
→ Tella / Opus Clip / Loom (con resumen de IA)
- Detectar oportunidades laborales según tu perfil tech**
→ Google Alerts + Reclaim + Notion Tracker

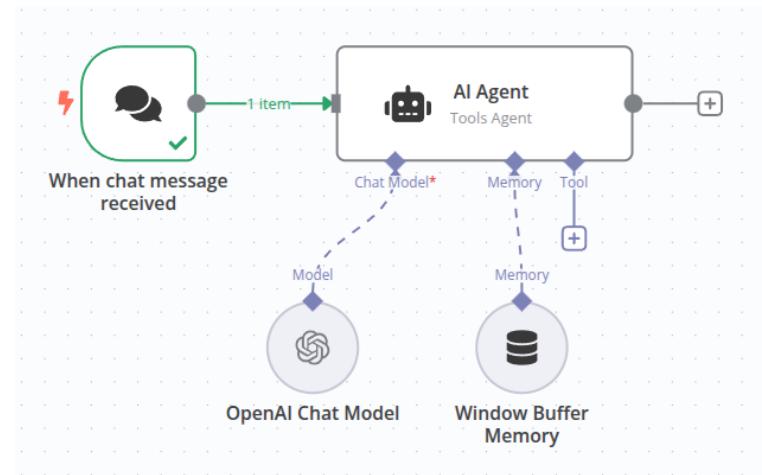
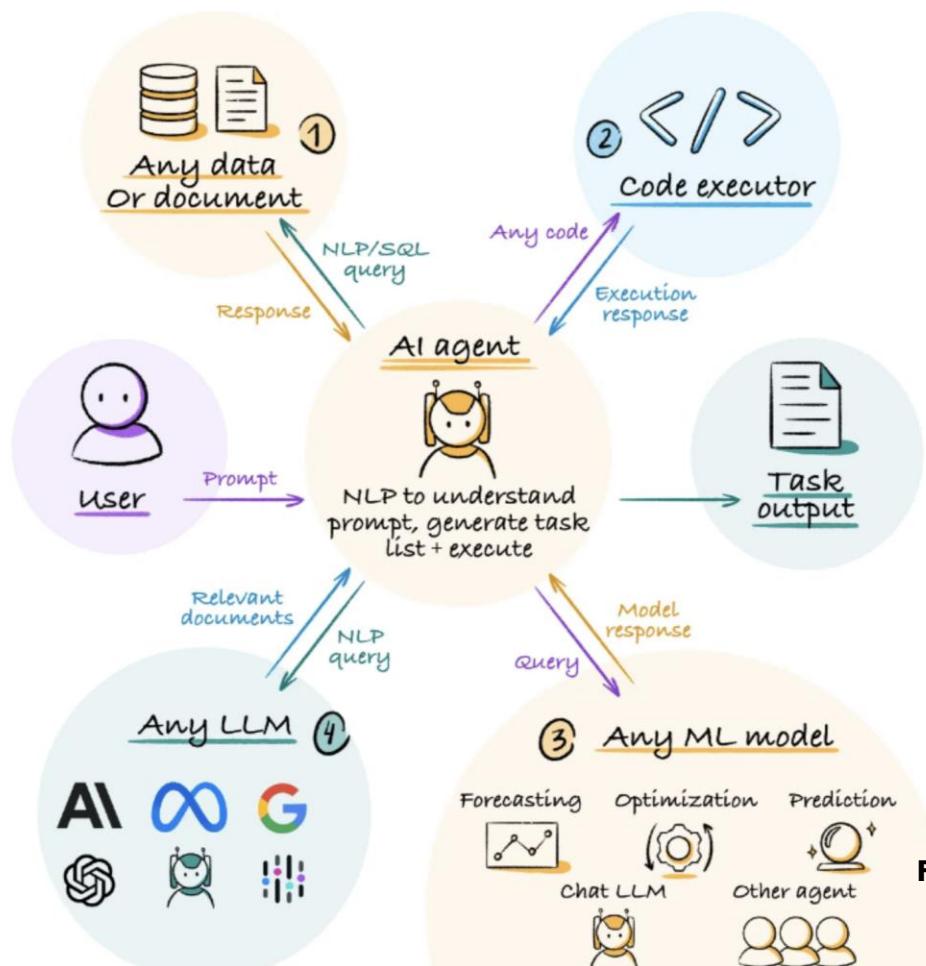
¿Para qué es clave automatizar hoy? Para:

- Trabajar con más foco en lo que requiere pensamiento crítico.
- Ahorrar horas en tareas repetitivas.
- Aprender habilidades con demanda real en el mercado Tech.

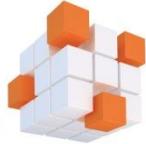


APLICACIONES – AGENTES IA

“Los Agentes IA son programas creados para percibir su entorno y tomar decisiones automáticas utilizando modelos de IA.”

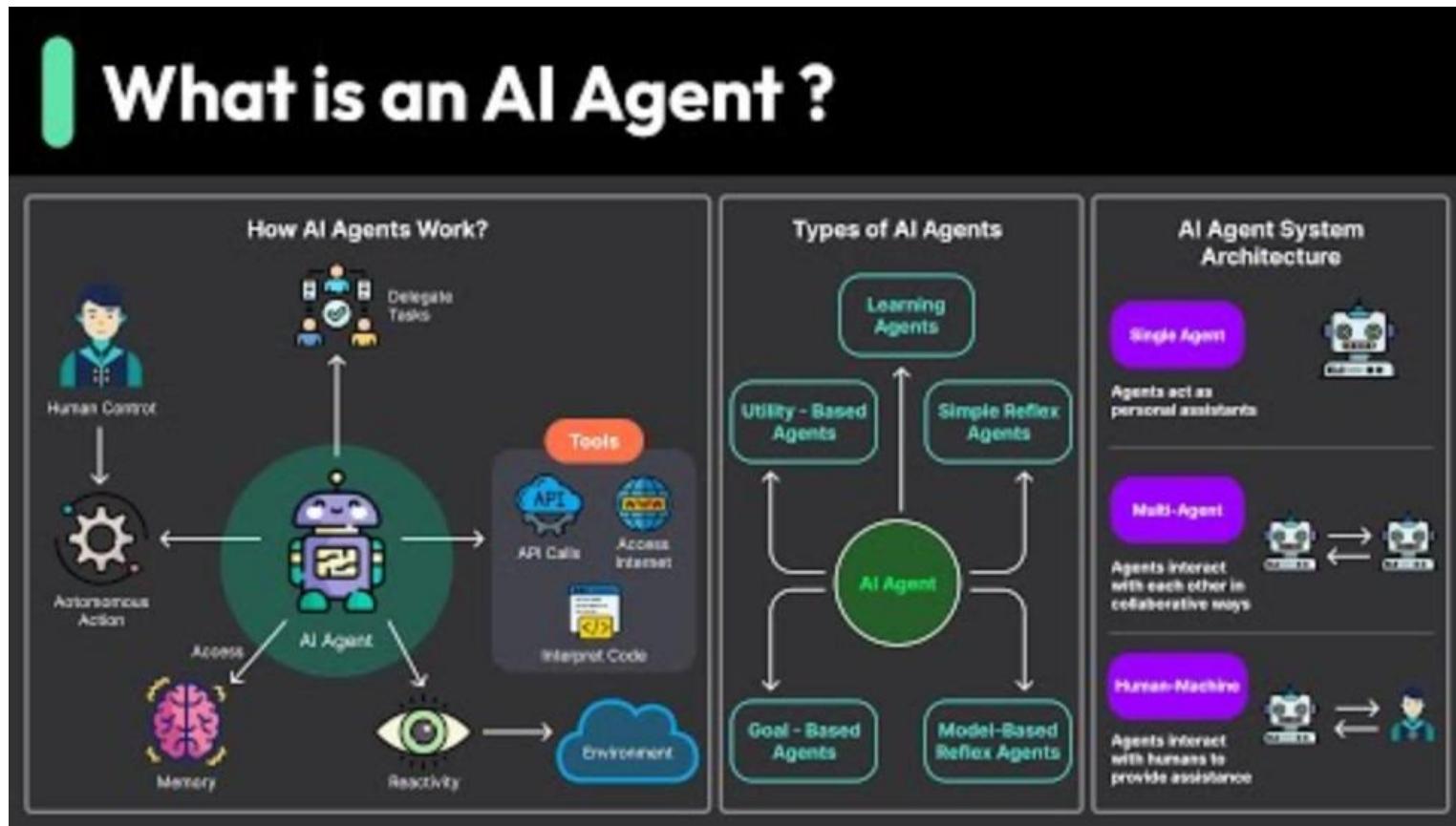


Fuente: [Learn AI Agents Handbook - Interactive Handbook for Building AI Agents | Learn AI Agents Handbook](#)
► [Anjana Data & LLM Hub - Anjana Data What Are AI Agents Really About?/](#)



APLICACIONES – AGENTES IA

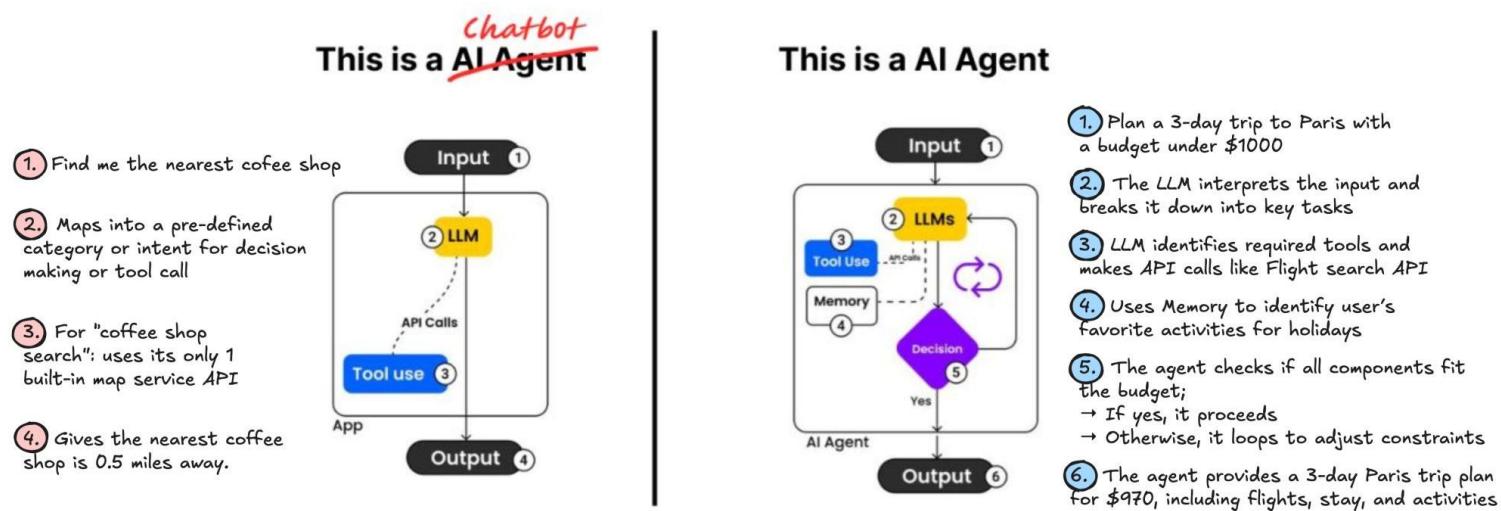
- **Video 1:** [What Are AI Agents Really About? - YouTube](#)



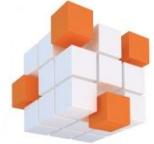


APLICACIONES – AGENTES IA

- Decir solo que “Un agente es solo un LLM que puede usar herramientas”, es lo mismo que decir que “es un chatbot sofisticado, no un agente”.
- Un verdadero agente de IA consta de:
 - Un LLM (para razonar)
 - Uso de herramientas (para actuar/usuarios herramientas)
 - Memoria (corta y larga)
 - Autorreflexión (para ajustarse según los resultados y el entorno)
 - Un agente no solo hace, sino que aprende, se adapta y decide qué hacer a continuación (de forma autónoma).
 - Habla con claridad.
 - Construye de forma más inteligente.
 - Llama a las cosas por su nombre.



Fuente: <https://www.linkedin.com/in/eordax/>
[AI Agents Roadmap - roadmap.sh](#)
[AI Engineer Roadmap](#)



DIFERENCIAS LLM VS AGENTE IA

- Un agente de IA se basa en grandes modelos de lenguaje (LLM).
- Mientras que los **LLM** solo procesan insumos para producir salida, **los agentes de IA** agregan funcionalidad orientada a objetivos.
- Pueden usar herramientas, procesar sus salidas, y **tomar decisiones** para completar tareas y resolver problemas.

Característica	LLM	Agente de IA
Capacidad principal	Generación de texto	Finalización de tareas orientadas a objetivos
Toma de decisiones	Ninguno	Sí
Usos Herramientas/APIs	No	Sí
Complejidad del flujo de trabajo	Un solo paso	Múltiples pasos
Alcance	Genera lenguaje	Realiza tareas complejas del mundo real
Ejemplo	LLM generando un párrafo	Un agente que programa una cita



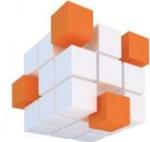
APLICACIONES CON IA

- **Video 1:** [Tutorial N8N: ¡Construir Chatbot en WhatsApp! \(Método fácil\) – YouTube](#)



- **Video 2:** [Cómo usar Claude para crear y replicar CUALQUIER Agente IA en n8n - YouTube](#)





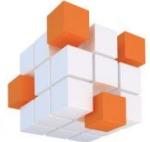
EXPLAINABLE AI (XAI)

¿Qué es XAI?

- **Explainable AI (XAI)** o **Explicabilidad en la Inteligencia Artificial** es un conjunto de métodos y técnicas que permiten a los humanos **comprender y confiar** en los resultados generados por modelos de inteligencia artificial (IA).
- Su objetivo es hacer que los modelos sean **transparentes, interpretables y auditables**, **evitando** el problema de las "**cajas negras**" donde los algoritmos producen resultados sin explicar cómo llegaron a ellos.

¿Por qué es importante?

- **Confianza:** Facilita la adopción de IA al brindar explicaciones claras.
- **Ética y Cumplimiento:** Asegura que los modelos sean justos y no discriminatorios.
- **Depuración de Modelos:** Ayuda a los científicos de datos a mejorar y depurar modelos.
- **Regulación:** Cumple con normativas como el GDPR, que exige explicabilidad en IA.



EXPLAINABLE AI (XAI)

Ejemplos de XAI en acción

- **Crédito Bancario:** Un modelo de IA predice si alguien es apto para un préstamo. Con XAI, se explica que la decisión se basa en el historial de pagos, ingresos y deudas previas.
- **Diagnóstico Médico:** Un modelo de IA detecta cáncer en imágenes médicas. XAI resalta las áreas críticas de la imagen que llevaron a la predicción.
- **Automóviles Autónomos:** Un coche autónomo frena repentinamente. XAI muestra que detectó un peatón cruzando la calle.
- **Detección de Fraude:** Un banco usa IA para detectar transacciones fraudulentas. XAI explica que se basó en patrones de gasto inusuales y ubicación.



EXPLAINABLE AI (XAI)

Herramientas prácticas para XAI

- **LIME (Local Interpretable Model-agnostic Explanations)**

Método basado en generar variaciones de la entrada para ver cómo cambia la predicción.

- **SHAP (SHapley Additive exPlanations)**

Explica la contribución de cada variable en la predicción del modelo.

- **InterpretML (Microsoft)**

Proporciona modelos explicables como Explainable Boosting Machine (EBM).

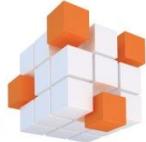
- **IBM AI Explainability 360**

Toolkit con diversas técnicas de explicabilidad para IA.

- **KNIME (No-Code XAI)**

KNIME ofrece nodos específicos para interpretar modelos de ML con técnicas como SHAP y LIME, sin necesidad de programar.

[Ver ejemplos en KNIME y ANEXO I](#)



APLICACIONES CON XAI

- **Video 1:** [Learning \(and teaching\) Explainable AI \(XAI\) with KNIME](#)

Two Distinct Approaches

- XAI: explaining black box models
- Interpretable ML: building transparent from the start



ÉTICA Y RESPONSABILIDAD EN EL USO DE LA MD/IA

MOTIVACIONES:

- Inteligencia Artificial está creciendo a gran velocidad.
- Mayores **incentivos** por tener **acceso** a la **información** → tanto en el sector privado como en el público.
- Proceso de recolección de datos → cada vez más **eficiente** y **económico** → es **urgente la necesidad de regular** este proceso.

¡LAS DECISIONES ALGORÍTMICAS AFECTAN PERSONAS!

PRINCIPIOS FUNDAMENTALES

- **Imparcialidad**: Tratar a todas las personas de manera justa, evitando sesgos por género, raza u otros factores.
- **Confiabilidad y seguridad**: Deben funcionar correctamente y sin fallas críticas (como en salud o vehículos autónomos).
- **Privacidad y protección de datos**: Los modelos usan datos personales; deben cuidarlos durante y después del entrenamiento.
- **Inclusión**: Que funcione para todos, sin excluir por discapacidad, identidad o contexto social.

Fuente: <https://deepnote.com/@mazzaroli/Etica-y-Manejo-de-Datos-para-Data-Science-e-Inteligencia-Artificial-6fbe98ba-5e1c-48ea-8b7b-42e98f6b2bac>



ÉTICA Y RESPONSABILIDAD EN EL USO DE LA MD/IA

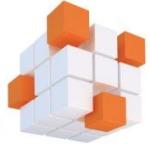
“Las aplicaciones con IA se ven humanas y generan confianza, pero pueden fallar y afectar personas si están mal entrenadas o diseñadas.”

Peligros si no se controla:

- Predicciones erróneas
- Discriminación por sesgos
- Pérdida de privacidad (datos sensibles)
- Poca transparencia

La responsabilidad recae en quienes desarrollan estas soluciones.

Los ingenieros deben anticipar riesgos, ser críticos y éticos.



ÉTICA, RESPONSABILIDAD Y SOSTENIBILIDAD EN EL USO DE LA IA

Ética en la IA

Preocupaciones:

- **Sesgo y discriminación:** La IA puede heredar y amplificar los sesgos presentes en los datos de entrenamiento, lo que lleva a decisiones injustas.
- **Privacidad y seguridad:** La IA puede recopilar y procesar grandes cantidades de datos personales, lo que plantea preocupaciones sobre la privacidad y la seguridad.
- **Impacto en el empleo:** La automatización impulsada por la IA puede provocar el desplazamiento de trabajadores en algunas industrias.
- **Armas autónomas:** El uso de la IA en sistemas de armas plantea cuestiones éticas sobre la responsabilidad y el control.

Marcos éticos:

- Principios y directrices para el desarrollo y uso de la IA, como los Principios de Asilomar (USA) y las directrices éticas de la UE.



ÉTICA, RESPONSABILIDAD Y SOSTENIBILIDAD EN EL USO DE LA IA

Uso Responsable de la IA

Principios de la IA Responsable:

- **Transparencia:** Ser claro sobre cómo funcionan los sistemas de IA y cómo toman decisiones.
- **Responsabilidad:** Establecer quién es responsable de las acciones de la IA y cómo se pueden abordar los errores.
- **Equidad:** Garantizar que la IA no discrimine ni perpetúe los sesgos.
- **Seguridad:** Proteger a las personas de daños causados por la IA.

Mejores prácticas:

- Estrategias para implementar la IA responsablemente, como realizar evaluaciones de impacto ético, involucrar a las partes interesadas y establecer mecanismos de supervisión.



ÉTICA, RESPONSABILIDAD Y SOSTENIBILIDAD EN EL USO DE LA IA

Uso Sostenible de la IA

Importancia para la IA:

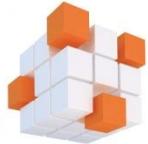
- Los datos de alta calidad son esenciales para entrenar modelos de IA precisos y confiables.
- La sostenibilidad de los datos garantiza que los datos sigan estando disponibles para su uso futuro en la IA.

Desafíos:

- Volumen de datos: La cantidad masiva de datos generados hoy en día plantea desafíos para el almacenamiento y el procesamiento.
- Consumo de energía: Los centros de datos consumen grandes cantidades de energía, lo que contribuye a las emisiones de carbono.

Estrategias:

- Eficiencia de datos: Reducir el desperdicio de datos mediante la recopilación y el almacenamiento solo de los datos necesarios.
- Energía renovable: Utilizar fuentes de energía sostenibles para alimentar los centros de datos.
- Eliminación responsable: Desechar los datos de una manera que proteja la privacidad y minimice el impacto ambiental.



ÉTICA Y RESPONSABILIDAD EN EL USO DE LA MD/IA



ÉTICA Y RESPONSABILIDAD EN EL USO DE LA MINERÍA DE DATOS E INTELIGENCIA ARTIFICIAL



Privacidad y protección de datos

La IA y el análisis de datos requieren grandes cantidades de información personal. Esto plantea preocupaciones sobre la privacidad de los individuos y la necesidad de proteger sus datos personales de un mal uso, como el robo, la manipulación o la divulgación no autorizada. Las regulaciones buscan establecer salvaguardias para garantizar que los datos se utilicen de manera ética y se protejan adecuadamente.

La necesidad de regular la inteligencia artificial (IA) y el uso de datos buscan equilibrar la innovación y los beneficios de la IA con la protección de los derechos y el bienestar de las personas, a través de:

Transparencia y explicabilidad

La falta de transparencia en cómo funcionan los sistemas de IA puede generar desconfianza y dificultar la comprensión de las decisiones tomadas por estos sistemas. Las regulaciones pueden exigir que los sistemas de IA sean transparentes y que se puedan explicar las razones detrás de sus decisiones. Esto es especialmente relevante en áreas críticas como la atención médica, donde la explicabilidad es fundamental para generar confianza en los diagnósticos y tratamientos.

Discriminación y sesgo

Los sistemas de IA pueden estar sujetos a sesgos inherentes debido a los datos utilizados para entrenarlos. Estos sesgos pueden resultar en decisiones discriminatorias o injustas en áreas como la contratación, la concesión de préstamos, la justicia penal, entre otros. La regulación busca mitigar estos sesgos y garantizar la equidad y la transparencia en las decisiones tomadas por los sistemas de IA.

Responsabilidad y rendición de cuentas

A medida que la IA se vuelve más autónoma, es necesario establecer responsabilidades claras en caso de daños causados por sistemas de IA. Las regulaciones pueden requerir que los desarrolladores y los operadores de IA sean responsables de las decisiones y acciones de los sistemas que crean y utilizan. Esto promueve la rendición de cuentas y asegura que haya mecanismos para abordar los posibles errores o daños causados por la IA.

Seguridad cibernética

A medida que la IA se vuelve más omnipresente, también aumenta el riesgo de ataques cibernéticos dirigidos a sistemas de IA y a los datos que utilizan. Las regulaciones pueden establecer requisitos de seguridad y estándares para proteger la integridad de los sistemas de IA y mitigar las vulnerabilidades que podrían ser explotadas por actores malintencionados.

ÉTICA Y RESPONSABILIDAD EN EL USO DE LA MINERÍA DE DATOS E INTELIGENCIA ARTIFICIAL



Autonomía y toma de decisiones

A medida que la IA se vuelve más autónoma, surge la cuestión de quién es responsable de las decisiones tomadas por los sistemas de IA. La regulación puede abordar la necesidad de establecer límites claros en la autonomía de la IA y definir la responsabilidad entre los desarrolladores, los operadores y los usuarios de estos sistemas.

Sustitución de empleo y desigualdad

La automatización impulsada por la IA tiene el potencial de sustituir trabajos humanos en ciertas industrias. Esto plantea preocupaciones sobre la desigualdad económica y social. Las regulaciones pueden buscar abordar este problema mediante la implementación de medidas de protección laboral, como la reeducación de los trabajadores o la creación de nuevos empleos relacionados con la IA.

Manipulación y desinformación

La IA puede ser utilizada para manipular y difundir información falsa o engañosa, lo que puede tener consecuencias perjudiciales para la sociedad y la democracia. La regulación puede abordar la necesidad de combatir la desinformación y garantizar la transparencia en la generación y difusión de contenido por parte de los sistemas de IA.

La necesidad de regular la inteligencia artificial (IA) y el uso de datos buscan equilibrar la innovación y los beneficios de la IA con la protección de los derechos y el bienestar de las personas, a través de:

Interacción humano-IA

A medida que la IA se integra más en nuestras vidas cotidianas, es importante considerar la interacción ética entre los seres humanos y los sistemas de IA. Las regulaciones pueden establecer pautas para la comunicación clara y ética entre los usuarios y los sistemas de IA, así como para evitar la dependencia excesiva o la explotación de los usuarios por parte de los sistemas de IA.

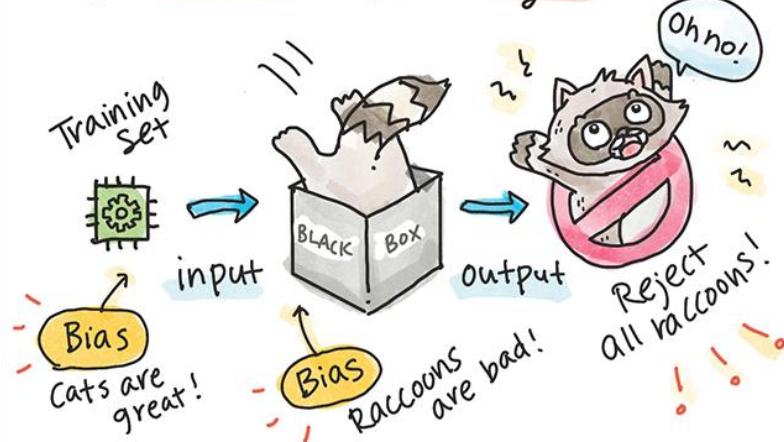
Impacto ambiental

El desarrollo y la implementación de la IA conllevan un consumo considerable de energía y recursos. La regulación puede buscar fomentar prácticas más sostenibles y responsables en términos de eficiencia energética y gestión de residuos, así como promover el desarrollo de soluciones de IA que ayuden a abordar los desafíos ambientales.

ÉTICA Y RESPONSABILIDAD EN EL USO DE LA MINERÍA DE DATOS E INTELIGENCIA ARTIFICIAL



Fairness in Machine Learning



Complex sociotechnical challenges



@Azure Advocates
@girlie_mac

Fairness-related harms

Unfairness = negative impacts for group of people such as those defined in terms of

- race • age
- gender • disability status

Harms:

- ★ Allocation
- ★ Quality of service
- ★ Stereotyping
- ★ Denigration
- ★ Over- / under- representation



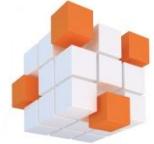
Assessment & mitigation

Fairlearn
fairlearn.github.io

- ♡ Identify the harm (+ benefits)
- ♡ Identify the affected groups
- ♡ Define fairness metrics

	False-Pos.	False-Neg.	Counts
men	0.35	0.21	6239
women	0.29	0.35	3124





ÉTICA, RESPONSABILIDAD Y SOSTENIBILIDAD EN EL USO DE LA IA

Desafío

Para **evitar que se introduzcan daños** en primer lugar, debemos:

- tener una **diversidad de antecedentes y perspectivas entre las personas que trabajan en sistemas.**
- invertir en **conjuntos de datos** que reflejen la **diversidad de nuestra sociedad.**
- desarrollar **mejores métodos** a lo largo del ciclo de vida del aprendizaje automático para **detectar y corregir** la IA **responsablemente** si hubiera un daño.



ÉTICA, RESPONSABILIDAD Y SOSTENIBILIDAD EN EL USO DE LA IA



El Futuro de la IA: Abrazar el Potencial con Responsabilidad

Puntos Clave:

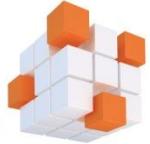
- La IA tiene el potencial de transformar muchos aspectos de nuestras vidas, pero es crucial abordar las consideraciones éticas y de sostenibilidad.

Perspectivas Futuras:

- El futuro de la IA depende de nuestra capacidad para desarrollar y utilizar esta tecnología de manera responsable.

Llamado a la Acción:

- Colaboración entre investigadores, formuladores de políticas, la industria y la sociedad para garantizar un futuro de la IA que sea beneficioso para todos.



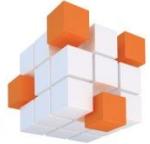
MATERIAL ADICIONAL CASOS PRÁCTICOS

LLMs y XAI en KNIME:

- [Intro to Large Language Models | KNIME](#)
- [GenAI with KNIME | KNIME](#)
- [How Banks Can Use Explainable AI in Credit Scoring | KNIME](#)
- [Explain Stroke Prediction Models with LIME in KNIME | KNIME](#)
- [Fundamentos del aprendizaje automático y la IA: producción de soluciones de IA explicable \(XAI\) y aprendizaje automático interpretable | LinkedIn Learning](#)

Otros Agentes IA (Dataiku, n8n, :

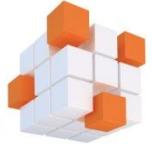
- [Dataiku Natural Language Processing \(NLP\) Recipes](#)
- [Project description - Wiki | Dataiku](#)
- [AI Agents With Dataiku | Dataiku](#)
- [Tutorial: Build an AI workflow in n8n | n8n Docs](#)
- [Learn AI Agents Handbook - Interactive Handbook for Building AI Agents | Learn AI Agents Handbook](#)
- [AI Agents Are Changing the World | Learn AI Agents Handbook](#)
- [AI Agents Engineer - roadmap.sh](#)



REFERENCIAS

Referencias:

- [Aplicaciones de IA para desarrolladores de .NET | Microsoft Learn](#)
- [Introducción a los conceptos de inteligencia artificial generativa - Training | Microsoft Learn](#)
- [Uso Responsable de la IA en la Educación](#)
- [6 tendencias de IA que verán más en 2025 - Source LATAM](#)
- [Announcing new tools in Azure AI to help you build more secure and trustworthy generative AI applications | Microsoft Azure Blog](#)
- [Caso de IA: Takeda | LinkedIn](#)
- [Driving Enterprise Transformation With Generative AI | Dataiku](#)
- [Prompt Engineering Roadmap - roadmap.sh](#)
- [AI Engineer Roadmap](#)
- [Interpretable Machine Learning](#)
- [Curso “IAG en el aula: herramientas para potenciar tus cursos” | Departamento de Derecho y Ciencia Política | UNLaM \(2025\)](#)
- Material libre uso publicaciones LinkedIn, Media Chicas, Jump Educación



MATERIAL ADICIONAL

- **Microsoft / ML-For-Beginners / Equidad**

<https://github.com/microsoft/ML-For-Beginners/blob/main/1-Introduction/3-fairness/README.md>

- <https://www.microsoft.com/en-us/ai/responsible-ai-resources?activetab=pivot1:primaryr4&rtc=1>

- <https://www.microsoft.com/en-us/research/theme/fate/>

- **Identificar los principios rectores para un manejo responsable de la inteligencia artificial**

https://learn.microsoft.com/es-es/training/modules/responsible-ai-principles/?WT.mc_id=academic-77952-leestott

- **Google AI**

<https://ai.google/responsibility/principles/>

<https://ai.google/responsibility/responsible-ai-practices/>

<https://developers.google.com/machine-learning/practica/fairness-indicators?hl=es-419>

- **Facebook (META) AI**

<https://ai.facebook.com/responsible-ai/>

- **IBM AI**

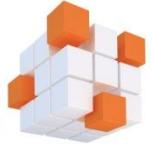
<https://www.ibm.com/mx-es/topics/ai-ethics>

<https://www.ibm.com/downloads/cas/J2LAYLOZ>

- **Sesgo Algorítmico**

<https://www.youtube.com/watch?v=nkOPXgpWS9w>

<https://www.youtube.com/watch?v=onJs6DXczu0>



MATERIAL ADICIONAL

- **UNESCO**

<https://news.un.org/es/story/2021/11/1500522>

<https://www.unesco.org/es/artificial-intelligence/recommendation-ethics#:~:text=Los%20Estados%20miembros%20de%20la,normativo%20mundial%20sobre%20el%20tema>

- **Future of Life Institute**

<https://futureoflife.org/>

- **HumanIA**

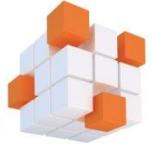
<https://www.chicos.net/humania/>

- **Ética**

<https://www.youtube.com/watch?v=1if8F7CKJc4>

Notas:

- [Agentes de IA: qué son y cómo cambiarán nuestra forma de trabajar - Source LATAM](#)
- [Elon Musk y otros líderes piden pausar la inteligencia artificial por los riesgos para la humanidad - Forbes Colombia](#)

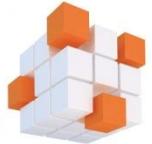


ANEXO I

EXPLAINABLE AI (XAI) EN KNIME - GUÍA PASO A PASO

En esta actividad práctica, exploraremos cómo interpretar modelos de Machine Learning usando KNIME, una plataforma de analítica de datos sin código. Utilizaremos LIME y SHAP, dos de las técnicas más populares en XAI.

-
- ◆ **Caso Práctico: Predicción de Churn con XAI en KNIME**
 - ◆ **Objetivo:** Explicar predicciones de un modelo de clasificación (Random Forest) usando SHAP y LIME.
 - ◆ **Datos:** Dataset de churn de clientes (puedes usar el dataset "Telco Customer Churn" de Kaggle).
 - ◆ **Herramientas:** KNIME Analytics Platform.
 - ◆ **Resultados:** Generar explicaciones visuales sobre qué variables afectan la predicción de abandono.
-



ANEXO I

EXPLAINABLE AI (XAI) EN KNIME - GUÍA PASO A PASO

❖ Paso a Paso en KNIME

1 Importar Datos

📌 Nodos a utilizar:

- CSV Reader → Cargar el dataset de churn.
- Missing Value → Manejo de valores nulos.
- Normalizer → Escalar los datos si es necesario.

2 Entrenar el Modelo de Machine Learning

📌 Nodos a utilizar:

- Partitioning → Dividir los datos en entrenamiento (80%) y prueba (20%).
- Random Forest Learner → Entrenar un modelo de clasificación.
- Random Forest Predictor → Aplicar el modelo a los datos de prueba.
- Scorer → Evaluar la precisión del modelo.

3 Interpretar el Modelo con LIME

📌 Nodos a utilizar:

- LIME for Classification (KNIME LIME Extension)
 - o Explica una única predicción generando datos similares y analizando los cambios en la predicción.
 - o LIME Table View
 - o Visualiza la importancia de las variables en una predicción específica.

📌 Flujo en KNIME:

• Conectar el Random Forest Predictor con LIME for Classification → Visualizar con LIME Table View.

✓ Ejemplo de interpretación con LIME:

Si un cliente es clasificado como "Churn", LIME puede indicar que la duración del contrato y el monto de la factura mensual fueron factores clave en la predicción.



ANEXO I

EXPLAINABLE AI (XAI) EN KNIME - GUÍA PASO A PASO

❖ Paso a Paso en KNIME

4 Interpretar el Modelo con SHAP

📌 Nodos a utilizar:

- SHAP Tree Explainer (KNIME SHAP Extension)
o Explica la contribución de cada variable a la predicción.
- SHAP Summary Plot
o Visualiza el impacto global de cada variable en el modelo.
- SHAP Dependence Plot
o Muestra la relación entre una variable específica y la predicción.

📌 Flujo en KNIME:

📍 Conectar el Random Forest Predictor con SHAP Tree Explainer → Luego con SHAP Summary Plot para análisis global o SHAP Dependence Plot para análisis individual.

Ejemplo de interpretación con SHAP:

Podemos ver que "Contract Length" tiene un impacto negativo en el churn (contratos más largos reducen la probabilidad de abandono).



ANEXO I

EXPLAINABLE AI (XAI) EN KNIME - GUÍA PASO A PASO

Actividad Final

- 1** Entrenar un modelo de Random Forest en KNIME para predecir Churn.
- 2** Usar LIME para interpretar una predicción específica y analizar qué factores influyeron en el resultado.
- 3** Usar SHAP para analizar la contribución de cada variable en el modelo completo.
- 4** Exportar los resultados e integrar un Dashboard en Power BI con los insights explicados.

Resumen Final

- ✓ LIME es útil para explicar decisiones individuales del modelo.
- ✓ SHAP permite interpretar la importancia global y local de cada variable.
- ✓ KNIME facilita la implementación de XAI sin necesidad de código.
- ✓ Power BI puede usarse para visualizar los resultados en un dashboard.

