

IACADEMY

DE FUNDAMENTOS A ARQUITECTURA DE IA

MÓDULO 01

Cómo funciona la IA

Fundamentos

iacedemy.com — 2026

MÓDULO 01

Cómo funciona la IA

Nivel: Fundamentos

Autor: Ricardo Gutierrez

Publicación: Mayo 2026

Plataforma: iacademy.com

Este material es parte del curso completo de IAcademy.

Uso personal e intransferible. Queda prohibida su redistribución o reproducción sin autorización.

Que es un LLM y por que te importa

Un LLM (Large Language Model) es un programa que ha leído miles de millones de textos y ha aprendido patrones estadísticos. Cuando le pides algo, **predice la siguiente palabra mas probable**. No piensa, no razona, no busca en una base de datos. Predice.

Si escribes "El cielo es...", el modelo dice "azul" porque en los textos que leyo, despues de esa frase casi siempre viene "azul". Esto funciona sorprendentemente bien para generar texto, código, traducciones y analisis.

Pero tiene consecuencias directas que necesitas entender:

1. **No tiene memoria real.** Cada conversacion empieza de cero (a menos que configures memoria persistente, como veremos en el M03).
2. **No verifica datos.** Puede inventar cifras, URLs, citas de libros que no existen. Se llaman "alucinaciones".
3. **La calidad depende de tu peticion.** Basura entra, basura sale. Un prompt vago produce una respuesta vaga.
4. **No ejecuta código por si solo.** Genera texto que parece código. Excepciones: Claude Code, que si puede ejecutar, y ChatGPT con Code Interpreter.

Regla de oro

Trata la IA como un **junior muy rapido**, no como un senior infalible. Siempre verifica datos criticos, siempre testea el código generado.

Que puede y que no puede hacer

Esta tabla te sirve como referencia rapida para saber cuando la IA te va a ayudar y cuando no:

Puede hacer bien

- Redactar textos, emails, documentos con estructura definida

- Analizar documentos largos y extraer informacion clave
- Generar codigo funcional (con revision humana)
- Traducir entre idiomas con calidad profesional
- Resumir, clasificar y organizar informacion
- Brainstorming y generacion de ideas
- Explicar conceptos complejos de forma simple

No puede hacer (o hace mal)

- Verificar si un dato es verdadero o falso
- Recordar conversaciones anteriores sin configuracion
- Acceder a internet en tiempo real (excepto Gemini y Perplexity)
- Garantizar que el codigo funciona sin testear
- Tomar decisiones eticas o legales por ti
- Contar con precision (contar palabras, caracteres, elementos)
- Matematicas complejas sin herramientas (calculadoras, Code Interpreter)

[IMG: Infografía de dos columnas: "La IA puede" (verde) vs "La IA no puede" (rojo) con iconos para cada punto]

Alucinaciones: el peligro silencioso

Las alucinaciones son el problema mas peligroso de la IA porque son invisibles. El modelo genera texto que suena perfecto, autoritativo, lleno de detalles. Pero es completamente inventado.

Ejemplos reales de alucinaciones:

- URLs que no existen pero tienen formato perfecto
- Citas de papers academicos con autores, año y revista. Todo falso.
- Estadísticas con decimales que parecen de un estudio real
- APIs y funciones de programación que no existen en la librería mencionada

Como prevenirlas: pide fuentes, verifica datos criticos, y anade en tu prompt: "Si no estas seguro de un dato, dilo explicitamente. No inventes."

Elegir entre ChatGPT, Claude, Gemini y DeepSeek

No uses la misma IA para todo. Cada una tiene un punto fuerte. Aquí tienes la comparativa real (no la del marketing):

ChatGPT (OpenAI, GPT-4o)

Mejor para: redacción creativa, brainstorming, contenido de marketing, generación de imágenes con DALL-E.

Punto débil: en documentos largos pierde contexto, tiende a ser verboso, a veces ignora restricciones del prompt.

Contexto: 128K tokens (~90 páginas).

Claude (Anthropic, Opus 4 / Sonnet 4)

Mejor para: análisis de documentos largos, seguir instrucciones complejas, código profesional, tareas con muchas restricciones.

Punto débil: más conservador en temas sensibles, puede ser demasiado cauteloso.

Contexto: Hasta 1M tokens (~700 páginas). El contexto más grande del mercado.

Gemini (Google)

Mejor para: búsqueda de información actualizada (conectado a Google Search), síntesis con fuentes frescas, integración con Google Workspace.

Punto débil: menos preciso en formato, a veces se salta restricciones del prompt.

Contexto: Hasta 1M tokens.

DeepSeek

Mejor para: código y razonamiento matemático, ejecución local (open-weight), soberanía de datos.

Punto débil: interfaz menos pulida, capacidad de escritura inferior a Claude/ChatGPT.

Contexto: 128K tokens.

Tabla de decision rapida

- **Redactar un email comercial:** ChatGPT
- **Analizar un contrato de 50 paginas:** Claude
- **Buscar datos de esta semana:** Gemini
- **Codigo y logica:** Claude Code o DeepSeek
- **Despliegue local (soberania):** DeepSeek o Qwen

Tu primera peticion profesional

La diferencia entre un resultado mediocre y uno excelente no es la IA que usas. Es **como le pides**.

Peticion mala

Resume este articulo.

Resultado: parrafo generico de 200 palabras sin estructura ni utilidad.

Peticion con estructura CORT

Contexto: Soy el director de marketing de una pyme de ciberseguridad con 50 empleados. Preparo un resumen semanal para el CEO.

Objetivo: Necesito un resumen ejecutivo de este articulo en 3 bullet points para incluirlo en el informe semanal.

Restricciones: Maximo 100 palabras total. Sin tecnicismos. Incluir las cifras clave del articulo. No repetir informacion.

Tono: Profesional, directo, sin relleno ni introducciones.

Mismo articulo, misma IA. El segundo resultado es 10 veces mas util porque le has dado contexto (quien eres), objetivo (que necesitas exactamente), restricciones (que NO quieres), y tono (como debe sonar).

[IMG: Captura de pantalla comparando los dos resultados en Claude. Izquierda: resultado generico. Derecha: resultado CORT con bullets concretos.]

La estructura CORT

CORT es el framework que usamos en todo el curso para estructurar peticiones:

- **Contexto:** Quien eres, que situacion tienes, para que proyecto trabajas.
- **Objetivo:** Que quieres que haga. Especifico, medible.
- **Restricciones:** Que NO debe hacer. Limites de longitud, formato, contenido.
- **Tono:** Como debe sonar. Formal, casual, tecnico, ejecutivo.

No necesitas las 4 siempre. Para tareas simples, Objetivo + Restricciones basta. Pero cuando el resultado no sea lo que esperas, anade mas componentes.

Diferencia entre modelos y cuando usar cada uno

Dentro de cada proveedor hay varios modelos. No es lo mismo usar GPT-4o que GPT-4o-mini, ni Claude Opus que Claude Haiku. Aqui tienes la guia practica:

Modelos grandes (Opus, GPT-4o, Gemini Ultra)

Mas caros, mas lentos, mas precisos. Para tareas complejas: analisis legal, arquitectura de software, documentos estrategicos. Cuando el coste de un error es alto.

Modelos medianos (Sonnet, GPT-4o-mini)

Balance entre coste y calidad. Para el 80% de las tareas diarias: redaccion, codigo, clasificacion, resumen.

Modelos pequenos (Haiku, GPT-3.5)

Muy rapidos, muy baratos. Para tareas de alto volumen y baja complejidad: clasificar 1000 emails, extraer datos de formularios, traducciones simples.

Regla del 80/20

El 80% de tus tareas se resuelve con un modelo mediano. Usa el grande solo cuando la precision es critica. Usa el pequeno para tareas repetitivas de alto volumen.

Modelos open-weight vs propietarios

Los modelos open-weight (DeepSeek, Qwen, Llama, Phi) se pueden descargar y ejecutar en tu servidor. Los propietarios (GPT, Claude, Gemini) solo se usan via API o interfaz web.

Cuando usar open-weight:

- Necesitas soberania de datos (no puedes enviar informacion a terceros)
- Tienes volumen alto y quieres reducir costes
- Necesitas personalizar el modelo (fine-tuning)

Cuando usar propietarios:

- Quieres la maxima calidad sin gestionar infraestructura
- Necesitas funciones avanzadas (vision, herramientas, agentes)
- El volumen es bajo y el coste de API es asumible

Ejercicio practico

Ejercicio M01: Compara 3 IAs con la misma tarea

1. Elige un articulo de tu sector (o usa uno del blog de IAcademy).
2. Escribe una peticion con estructura CORT.
3. Enviala a Claude, ChatGPT y Gemini.
4. Compara los resultados en una tabla con estas columnas: IA, Calidad, Formato, Datos inventados, Nota (1-10).
5. Decide cual usarias para esa tarea y justifica por que.

Bonus: Crea un documento "AI Decision Guide" de 1 pagina para tu equipo. Que IAs usais, para que, y que reglas seguir. Usa la plantilla en los recursos del modulo.

Conclusiones clave

Key takeaways del M01

1. Un LLM predice palabras, no piensa. Siempre verifica datos criticos.
2. Las alucinaciones son el peligro mas grande porque son invisibles. Pide fuentes y confirma.
3. Cada IA tiene un punto fuerte: ChatGPT (creativo), Claude (precision), Gemini (actualizado), DeepSeek (local).
4. La estructura CORT (Contexto, Objetivo, Restricciones, Tono) transforma cualquier peticion de mediocre a profesional.
5. El 80% de tus tareas se resuelve con un modelo mediano. Reserva el grande para decisiones criticas.
6. Open-weight para soberania y volumen. Propietarios para maxima calidad sin infraestructura.

Siguiente: M02 - Prompting que funciona

Ahora que entiendes como funciona la IA, el siguiente paso es dominar los prompts: 7 componentes, 4 tipos, few-shot learning y debug.

[Ir al Modulo 02](#)

IACADEMY

iacedemy.com

De fundamentos a arquitectura de IA.
12 módulos prácticos. 24 recursos descargables.
Quizzes con certificado. Vídeos profesionales.

Empieza gratis en iacedemy.com/free

© 2026 IAcademy — Todos los derechos reservados