La IA generativa y nuestra docencia

Cuánto ha mejorado, a qué ritmo lo hace y cómo responden las universidades de referencia

Taller sobre IA generativa y metodología docente
Facultad de Economía y Empresa · Universidad de Murcia · julio de 2026

Dos formas de plantear el problema

La pregunta de 2023

«¿Cómo evito que mis estudiantes usen la IA?»

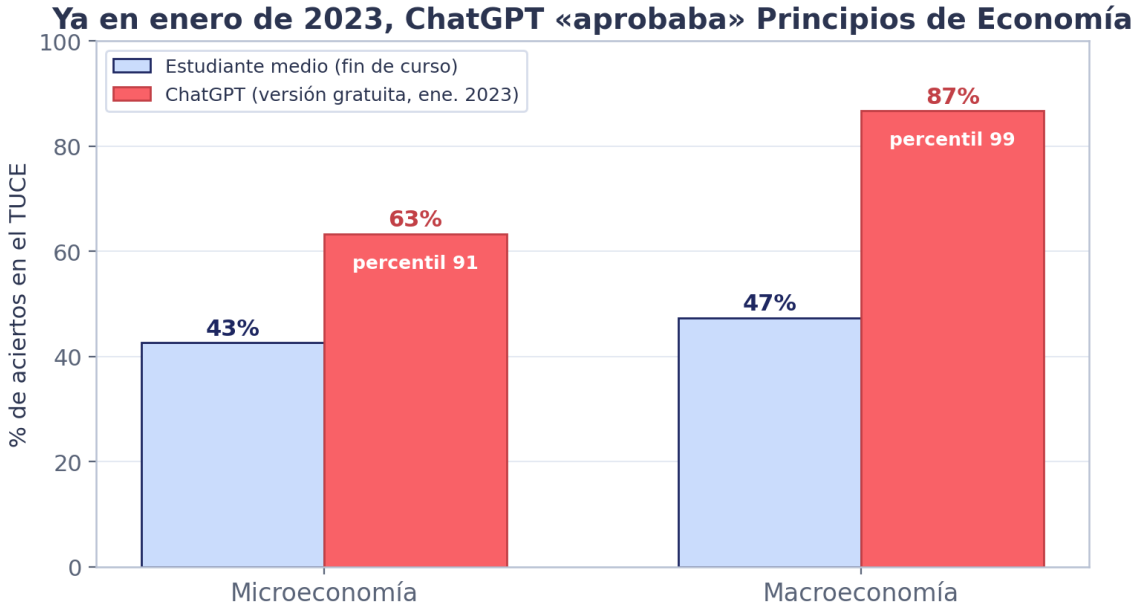
La pregunta de 2026

«¿Qué evidencia de aprendizaje me queda, y cómo la rediseño?»

95% de los estudiantes de grado ya usa IA en sus estudios; **94%** en trabajos evaluables (HEPI/Kortext 2026, Reino Unido)

Enero de 2023: el aviso llegó pronto (y a Economía)

ChatGPT gratuito frente al estudiante medio en el test estandarizado de principios (TUCE)

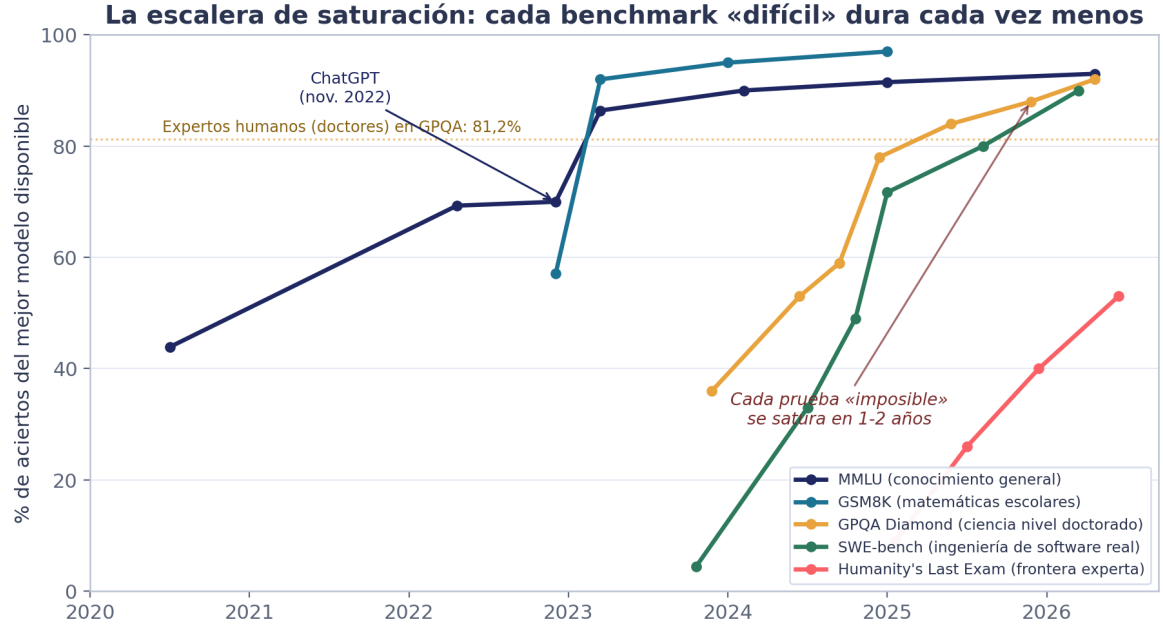


Fuente: Geerling, Mateer, Wooten y Damodaran (2023), The American Economist. Test of Understanding in College Economics (TUCE), muestra normativa de EE. UU.

Con la primera versión, ya percentil 91 en Micro y 99 en Macro.

Cada prueba «imposible» se satura en uno o dos años

Mejor resultado público en benchmarks sucesivamente más difíciles, 2020-2026

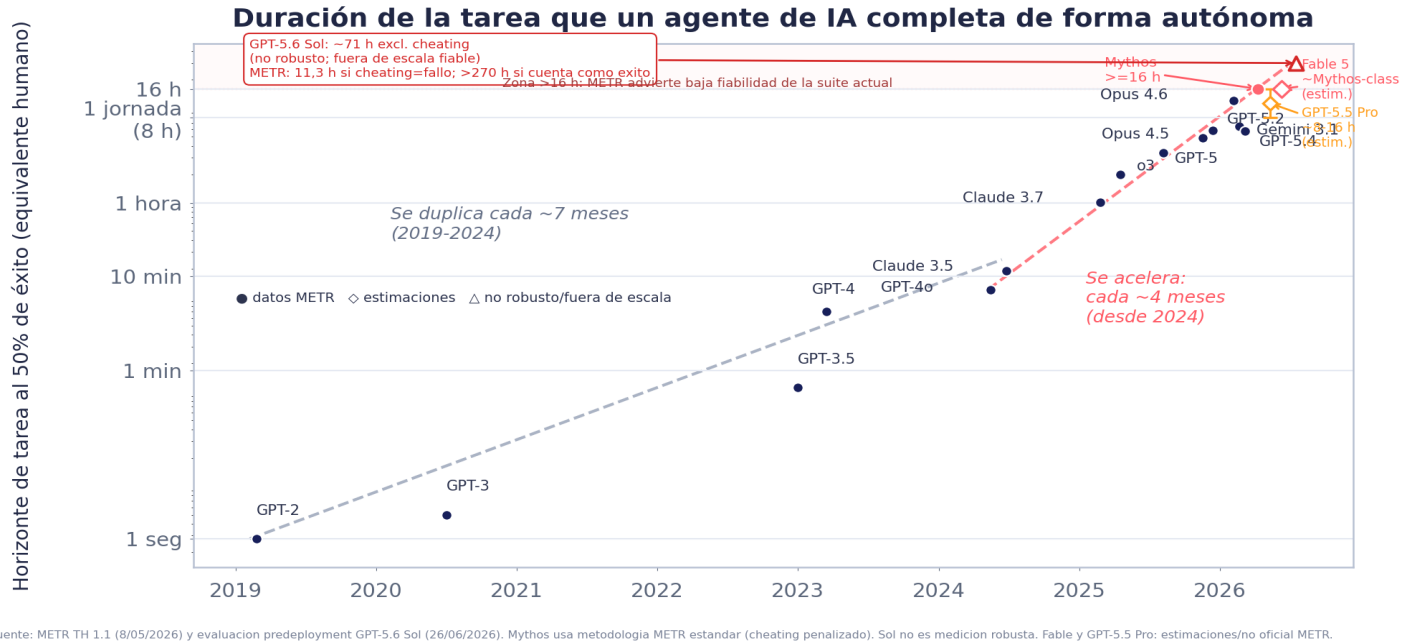


Valores aproximados del mejor modelo público en cada fecha. Fuentes: informes técnicos (GPT-4, 2023), Stanford AI Index 2025 y 2026, leaderboards públicos (Scale/CAIS, Artificial Analysis).

Los modelos ya superan a los expertos con doctorado en su propio terreno (GPQA).

De responder preguntas a completar trabajos

Duración de la tarea que un agente completa de forma autónoma (METR, escala logarítmica)



Fuente: METR TH 1.1 (8/05/2026) y evaluacion predeployment GPT-5.6 Sol (26/06/2026). Mythos usa metodologia METR estandar (cheating penalizado). Sol no es medicion robusta. Fable y GPT-5.5 Pro: estimaciones/no oficial METR.

El horizonte se duplica cada 4-7 meses: de minutos (2023) a jornadas completas (2026).

La prueba de fuego: la ciencia de frontera

De resolver problemas con solución conocida a producir conocimiento nuevo verificado

2025

Papers con revisión por pares

Primer artículo generado íntegramente por IA aceptado en un workshop de ICLR (Sakana, descrito en Nature). Meses después, otro entra en el programa principal de ACL: top 8,2% de los envíos.

Jul. 2025

Oro olímpico

Medalla de oro en la Olimpiada Internacional de Matemáticas: sistemas de razonamiento de Google DeepMind y OpenAI, al nivel de los mejores estudiantes del mundo.

May. 2026

Problema abierto desde 1946

Un modelo de OpenAI refuta la conjetura de la distancia unitaria de Erdős. Demostración verificada por Gowers, Sawin y otros: «se publicaría en una revista top si fuera humana».

Nuestras evaluaciones quedaron atrás hace tiempo: la frontera está ya en la investigación.

«La IA no sirve para esto»

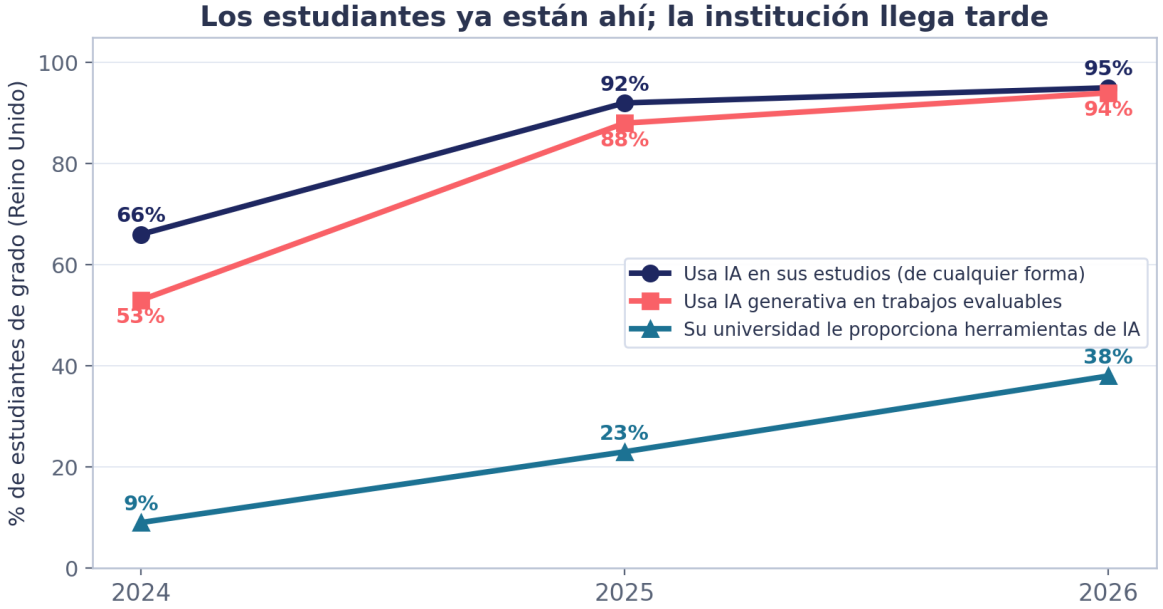
es una foto fija en medio de una película acelerada

Regla práctica: si una actividad no presencial resiste hoy «por poco», no resistirá el curso que viene. Diseñemos para la tendencia, no para la versión de hoy.

Cautela honesta: la frontera es «dentada» (Dell'Acqua et al., 2023) — brillante en tareas difíciles, torpe en variantes triviales. Verificar sigue siendo imprescindible. Y eso, precisamente, se enseña.

Mientras tanto, los estudiantes ya están ahí

Estudiantes de grado, Reino Unido (HEPI/Kortext 2024-2026)



Fuente: HEPI/Kortext, Student Generative AI Survey 2024, 2025 y 2026 (encuestas Savanta a más de 1.000 estudiantes de grado a tiempo completo en el Reino Unido).

La pregunta no es si lo permitimos: es qué usos ordenamos, declaramos y evaluamos.

Qué hacen las universidades de referencia

El mismo patrón en 2025-2026: de prohibir a estructurar

1 · Acceso

IA como infraestructura de campus: ChatGPT Edu en Cal State (460.000 estudiantes), Oxford, ASU, Bocconi… Claude for Education en LSE, Northeastern…

2 · Gobernanza

Reglas comunes (principios del Russell Group, matrices de uso por asignatura) y adiós a la detección como estrategia: un detector nunca es prueba única.

3 · Formación

Alfabetización en IA para estudiantes y profesorado: guías, talleres, comunidades de práctica internas (este taller).

4 · Rediseño

Evaluación, metodología y contenidos se revisan a la vez. Es el componente difícil — y el de las dos próximas diapositivas.

Evaluación: el modelo de dos carriles

Universidad de Sídney (2024-2025), el marco de referencia internacional

Un mismo título, dos tipos de evidencia

CARRIL 1 · SEGURO

Evaluación DEL aprendizaje

Presencial y supervisado:
exámenes, defensas orales,
prácticas, laboratorio

Uso de IA controlado o prohibido

Valida los resultados
del título

CARRIL 2 · ABIERTO

Evaluación PARA el aprendizaje

Trabajos, casos, proyectos,
ensayos, análisis de datos

Uso de IA permitido o exigido,
siempre declarado

Enseña a trabajar (bien)
con las herramientas

*«Cualquier entrega no supervisada puede completarse a buen nivel con IA;
prohibirla no es práctico, ni verificable, ni deseable» — A. Bridgeman (Univ. de Sídney)*

Adaptado del «two-lane approach» de la Universidad de Sídney (Liu y Bridgeman, 2024-2025), aplicado a todas sus titulaciones desde 2025.

Y vuelve la palabra: Cornell y Penn reintroducen la defensa oral de forma sistemática.

La IA dentro del aula: metodología y contenidos

Tutor y simulador	Marco de Wharton (Mollick): la IA como tutor, entrenador, mentor, compañero, herramienta, simulador o «estudiante» al que enseñar — cada papel con instrucciones y riesgos propios.
Copiloto docente	Harvard Business School: ChatLTV, entrenado con los casos y notas del curso, responde dudas 24/7 y libera la tutoría humana para lo que más valor añade.
Tarea de segunda vuelta	No «redacta un análisis», sino «genera dos análisis con IA, contrasta los datos y defiende cuál es más convincente». El valor se desplaza a la auditoría y la argumentación.
Subir el techo del temario	Como la calculadora o el ordenador en estadística: encargos «imposibles hasta ahora» (Wharton) y métodos avanzados que bajan a grado — con el suelo verificado en el carril seguro.

La misma herramienta suma o resta: depende del diseño

Experimento de campo, ~1.000 estudiantes (Bastani et al., Wharton, 2024)

Chat libre para «hacer los deberes»

−17%

en el examen posterior, cuando se retira la herramienta

Tutor con andamiaje pedagógico

Daño ≈ 0

pistas en lugar de soluciones: se conserva el esfuerzo que produce aprendizaje

La variable decisiva no es la herramienta: es el diseño de la tarea.

Y en para nosotros: cuatro movimientos realistas

1

Reglas explícitas por asignatura

Matriz simple presentando la asigntura: usos prohibidos · permitidos con declaración · recomendados · evaluados.

2

Revisar primero lo más expuesto

Las actividades no presenciales con más peso en la nota: son tareas «de horas», el rango de los agentes.

3

Proceso, oralidad y aplicación situada

Minidefensas con muestreo, hitos de proceso, datos reales, casos: evidencia que la IA no puede suplantar.

4

Comunidad de práctica interna

Compartir experiencias, plantillas y rúbricas, como este taller.

De la preocupación individual a la respuesta colectiva

Tres preguntas para el taller de hoy:

- ¿Qué actividades de mi asignatura ya no aportan evidencia de aprendizaje?
- ¿Qué haría en mi asignatura un carril seguro y un carril abierto?
- ¿Qué necesito de la Facultad para probarlo el curso que viene?

Gracias — seguimos en la parte práctica del taller