

predecir la próxima palabra → “Máquinas que **hablan**”

Máquinas que hablan

Qué es la IA generativa y por qué prestarle atención

ENTRE LA PROMESA Y EL RIESGO

Belén Saavedra

MSc · Université de Lorraine & Universidad del País Vasco · LCT Erasmus Mundus

Ateneo de Tesis · UNLu · 4 de julio

UN JUEGO · COMPLETÁ CON LA PRIMERA PALABRA

Buenos días,
¿cómo

p(próxima palabra)

¿qué palabra sigue?

UN JUEGO · COMPLETÁ CON LA PRIMERA PALABRA

Buenos días,
¿cómo estás?

p(próxima palabra)

 estás

 sos

 pasó

una palabra domina la distribución

UN JUEGO · COMPLETÁ CON LA PRIMERA PALABRA

Lo que más me preocupa
de la tesis es

p(próxima palabra)

¿qué palabra sigue?

UN JUEGO · COMPLETÁ CON LA PRIMERA PALABRA

Lo que más me preocupa
de la tesis es el tiempo

p(próxima palabra)

██████████ el tiempo

██████████ la bibliografía

██████████ la metodología

██████████ mi director

ahora la probabilidad se dispersa

ACTO 0 · ¿DE QUE HABLAMOS?

Lo mismo que ustedes — a escala

USTEDES · LA SALA

¿cómo estás?

...la tesis es **el tiempo**

=

EL MODELO

¿cómo estás?

...la tesis es **el tiempo**

Las mismas predicciones — pero expuesto a volúmenes enormes de texto. De tanto leer, predice con naturalidad. Por eso suena tan bien.

- UN CASO REAL

¿Y si lo probable no es lo verdadero?

EL PROMPT – ALGO ASÍ

"Pasame tres referencias académicas sobre el filtro afectivo en la inmersión bilingüe tardía, con autor, año y revista."

Devuelve tres citas de aspecto impecable.

Ninguna de las tres existe.

NO EXISTE

Krashen, S. D. (1987). *Affective filtering in late bilingual immersion*. Journal of Second Language Acquisition, 14(3), 221–245.

NO EXISTE

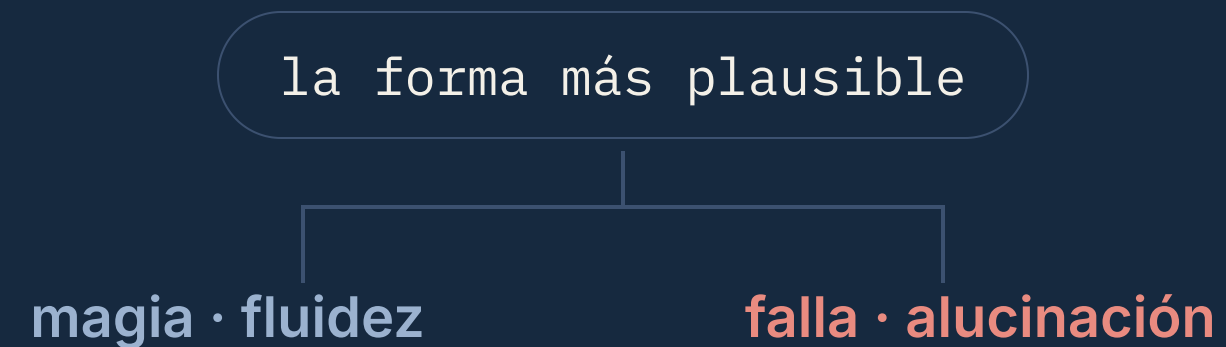
Long, M. H., & Doughty, C. (2009). *Negotiated interaction and uptake in task-based CALL*. Applied Linguistics Review, 7(2), 88–112.

NO EXISTE

Norton, B. (2014). *Investment, identity and the multilingual writing self*. TESOL Quarterly, 48(4), 510–537.

Fluidez y error salen del mismo mecanismo.

Un modelo predice la forma más plausible del texto. Eso es todo lo que vamos a desarmar hoy.



Primero, ¿de qué hablamos?

IA

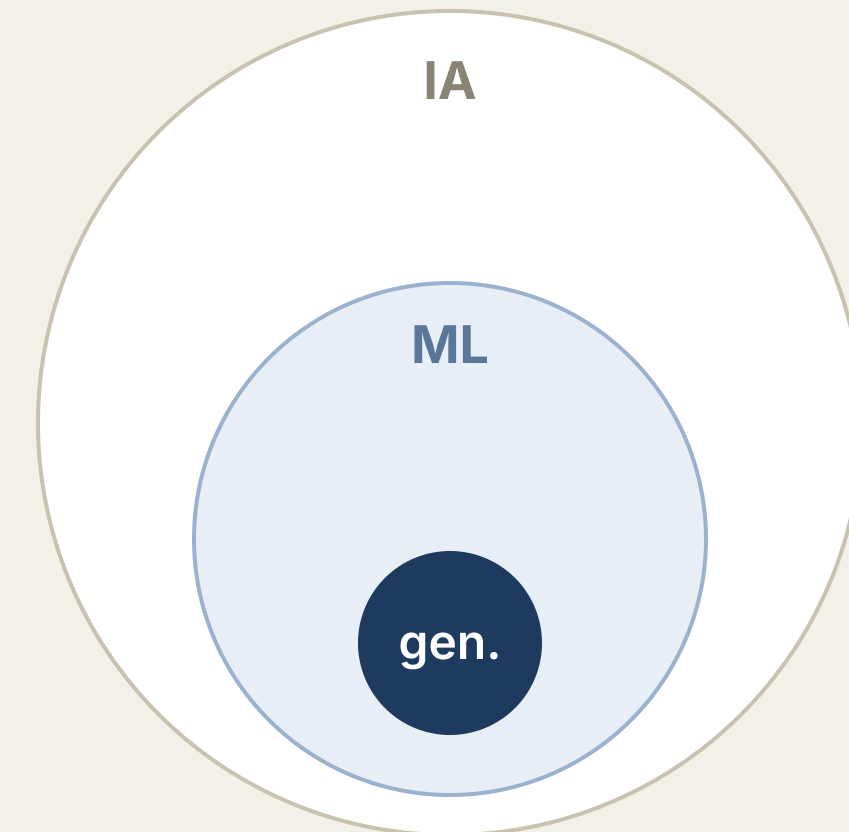
Sistemas que hacen tareas que asociamos con la inteligencia.

Machine learning

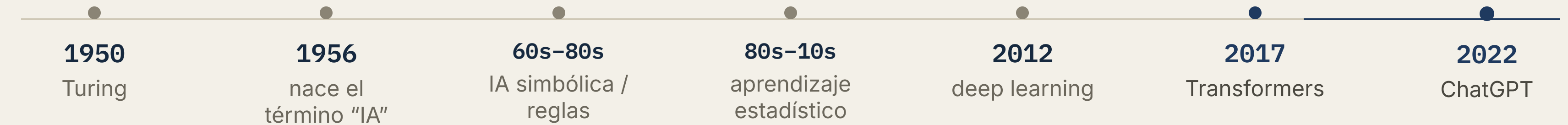
No se programa con reglas; aprende patrones de datos.

IA generativa

Produce contenido nuevo. Un punto chico y reciente.



No nació en 2022



Casi 70 años de IA que **clasifica, predice, optimiza, traduce.**

La generativa es el último tramo — una rama joven.

ACTO 1 · QUÉ ES Y CÓMO FUNCIONA

Un mapa de lo que hace la IA

La generativa es **un barrio del mapa** — el grupo de tareas donde predecir la forma más plausible es, justo, lo que querés.

APRENDIZAJE SUPERVISADO

Clasificar

elegir entre categorías dadas

spam

diagnóstico por imagen

sentimiento

MODELOS PREDICTIVOS

Predecir

estimar un valor futuro

demanda

riesgo crediticio

clima

VISIÓN · RECONOCIMIENTO DE PATRONES

Reconocer

identificar qué hay en una señal

voz

rostros

objetos

SISTEMAS DE RECOMENDACIÓN · OPTIMIZACIÓN

Ordenar & optimizar

priorizar o decidir entre muchas opciones

buscadores

feeds

rutas / logística

ACÁ ESTAMOS HOY

Generar IA GENERATIVA

producir una secuencia nueva, token a token — predecir la forma más plausible

redactar

resumir

traducir

escribir código

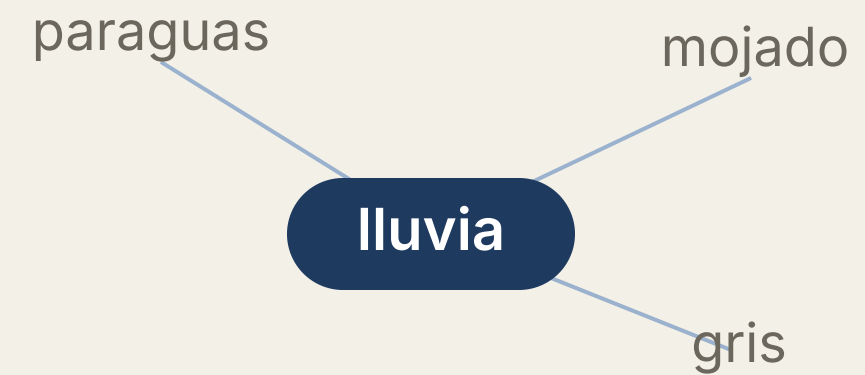
generar imágenes

Aprende la forma, no el significado

“Conocés una palabra por la compañía que mantiene.”

— Firth, 1957

El modelo aprende la **compañía estadística** de cada forma.
Nunca tocó el referente.



El pulpo en el cable



1 Escucha la forma

Intercepta millones de mensajes entre A y B. Aprende qué palabra suele seguir a cuál.

2 Se hace pasar

Predice tan bien que puede responderle a B haciéndose pasar por A. Nadie nota la diferencia.

3 Falla en lo real

B: «me persigue un oso, ¿qué hago?». El pulpo nunca vio el mundo — no puede ayudar.

Loros estocásticos

sin mente detrás



«Según los datos, es plausible que en este contexto la evidencia sugiera... cabe concluir que...»

suenan impecables ✓

Estocástico

Azar gobernado por probabilidades: elige cada palabra por lo que suele venir después.

Loro

Repite la forma del lenguaje sin conectarla con ningún significado.

La trampa es nuestra

Asumimos que detrás de un lenguaje fluido hay una mente que entiende. No la hay.

La misma habilidad, las dos caras



Donde plausible $\approx$ útil

Cuando la tarea es darle forma a un contenido, “lo más plausible” **es** el buen resultado.

Reformular

Resumir

Cambiar de registro

Ayuda en traducción

Estructurar

Pulir

Generar variantes

Desbloquear la página en blanco

Donde plausible $\neq$ verdadero

Alucinación

Hechos y citas inventados con formato perfecto.

EJ. «Krashen (1987), *Journal of SLA* 14(3)» — la revista nunca existió.

NO EXISTE

Seguridad sin verdad

Tono confiado, igual de errado.

EJ. «Saussure lo sostiene en el *Curso*»... y jamás lo escribió.

Sesgo

Amplifica lo dominante, borra lo minoritario.

EJ. Pedís un dibujo de «ingeniero» y describe a un hombre europeo por defecto.

Homogeneización

Aplana la voz y la variación lingüística.

EJ. El voseo o el guaraní se «corrigen» a un español neutro de manual.

Una regla práctica

✓ SÍ DELEGAR — la forma

Dar forma a algo que ya pensaste

Reformular, resumir, traducir

Estructurar, pulir, variar

Romper el bloqueo inicial

✗ NO DELEGAR — la cognición y la verdad

Pensar o argumentar por vos

Fuentes, datos, citas (sin verificar)

Decidir qué es verdad

Producir lo que después no podés defender

El orden es la variable



No prohibimos nada. Distinguimos: **ejecución (sí)** vs. **cognición (no)**.

Un compañero de razonamiento (con asterisco)

LO QUE SÍ ACELERA

Te explica a tu nivel (**las veces que sea necesario**)

Te da feedback inmediato

Juega al abogado del diablo

Presiona tus ideas

EL ASTERISCO

No razona: lo *simula*.

Le da forma a **tus** ideas; nunca las tiene por vos. El espejismo sirve solo si vos ponés el razonamiento real.

Deuda cognitiva

"Your Brain on ChatGPT", MIT (2025) · EEG · 54 personas

LA PREGUNTA DEL ESTUDIO

¿Qué le pasa a tu cerebro cuando le delegás la escritura a la IA? Lo midieron con EEG mientras la gente escribía ensayos, sesión tras sesión.

GRUPO 1 · DELEGÓ EN LA IA

Escribió siempre con ChatGPT.

↓ menos actividad cerebral y menos sentido de propiedad — muchos no podían citar una línea de su propio texto.

GRUPO 2 · PENSÓ PRIMERO

Escribió sin ayuda; sumó la IA recién después.

↑ mantuvo la actividad y la memoria; al sumar la IA, hasta rindió mejor.

EN AMBOS → la IA acható la voz: los textos se volvieron más parecidos entre sí — homogeneidad lingüística.



PARA CONTAR

Y esto se ve en la vida real

Nadie afirma que "te pudre el cerebro" — ni los propios autores. Es un preprint chico (n = 54, sin peer review), pero el efecto se percibió.

En enseñanza de lenguas, la producción ES la habilidad

Piénsenlo desde su aula: ustedes no enseñan a *mostrar* una L2, enseñan a *producirla* — y producir es, justamente, donde se aprende.

Si un alumno terceriza escribir en L2, no escribe peor: **no adquiere la lengua.**

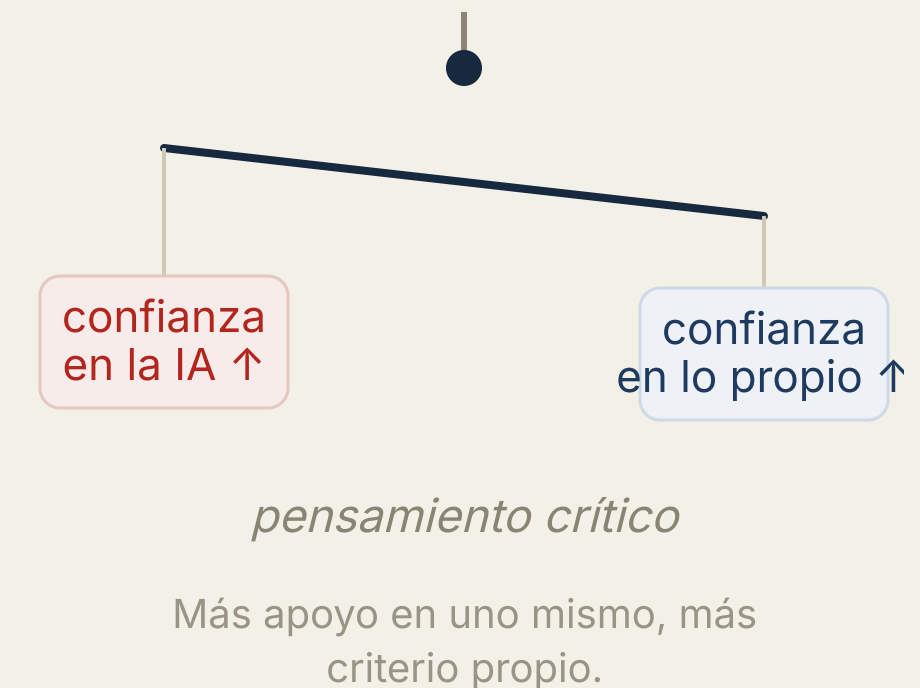
Swain · hipótesis del output

Bjork · dificultad deseable

EL ESTUDIO · MICROSOFT / CMU · 2025

Encuestaron a cientos de profesionales que usan IA a diario. Cuanta **más confianza ponen en la IA**, menos pensamiento crítico ejercen; cuanta **más confían en su propia competencia**, más.

El antídoto es la *competencia*: acelera a quien ya sabe, limita al que recién aprende.



¿Y AHORA?

~~¿Usar o no usar IA?~~

Ésa nunca fue la pregunta.

La pregunta real es **con qué herramientas** — y **en qué orden**.

Cosas concretas para llevarse →

Pedirle bien

- 1 Dale rol y contexto.
- 2 Trabajá sobre **tu material**: pegá tu texto, que no invente.
- 3 Pedí **proceso**: "criticá mi argumento", "dame el contraargumento más fuerte".
- 4 Que pregunte antes de responder y marque lo que no sabe.
- 5 Iterá: la 1ª respuesta es borrador, no veredicto.
- 6 Separá **forma** (delegá) de **hecho** (verificá).

PROMPT POBRE

"Escribime la introducción de mi tesis."

PROMPT CON CRITERIO

*"Sos mi tutor crítico. Acá va mi argumento [pego].
Criticalo, dame el contraargumento más fuerte y
preguntá lo que falte."*

Caja de herramientas, por fase

FASE	HERRAMIENTA	POR QUÉ
Descubrir literatura real	Semantic Scholar · Research Rabbit	papers reales, no inventados
Preguntar a la evidencia	Consensus · Elicit	respuestas con respaldo, citas a nivel oración
Verificar citas	Scite	si trabajos posteriores apoyan o contradicen
Sintetizar tus fuentes	NotebookLM	anclado en tus PDFs, cita la página, casi no alucina
Gestionar bibliografía	Zotero	orden y formato
Reformular · estructurar · criticar	Claude · ChatGPT	forma, NO hechos — verificá siempre

 ancladas en fuentes  modelos generales — solo forma

El protocolo, en 4 pasos

1

Tu argumento primero

La tesis es tuya.

2

IA para la ejecución

Forma, estructura, pulido.

3

Verificación

Toda cita y todo dato,
contra la fuente.

~~Krashen, S. D. (1987). Affective
filtering...~~

4

Apropiación

Si no lo podés defender, no
es tuyo todavía.

“...mantené al humano en el lugar del **humano**” ← la última predicción :)

La misma herramienta, dos caras.

BIEN USADA

Ayuda a **aprender**, a **razonar** y **acelera** el trabajo.

MAL USADA

Limita al que aprende, invalida los resultados y multiplica el **error a gran escala**.

Nuestra tarea es proteger la **competencia** y el **pensamiento crítico**, y mantener al **humano en el lugar del humano**.

Gracias

CONTACTO

Belén Saavedra

belen.saavedra.39@gmail.com

RECURSOS DE ESTA CHARLA

Kosmyna et al. (2025), *Your Brain on ChatGPT* · MIT ·
[arXiv:2506.08872](https://arxiv.org/abs/2506.08872)

Lee et al. (2025), *Generative AI and Critical Thinking* · Microsoft / CMU
· CHI 2025

Caja de herramientas: NotebookLM · Semantic Scholar · Elicit ·
Consensus · Scite · Zotero