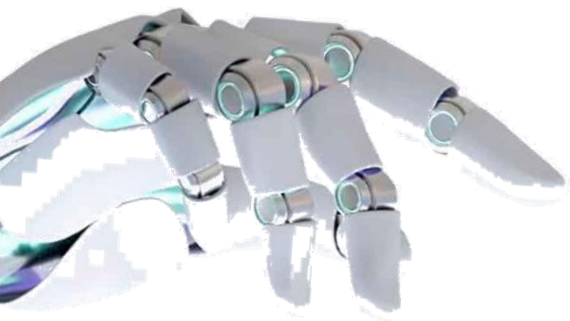


IA en la Investigación académica



Wilfrido A. Gómez Arias
wilfrido.gomez@ibero.mx

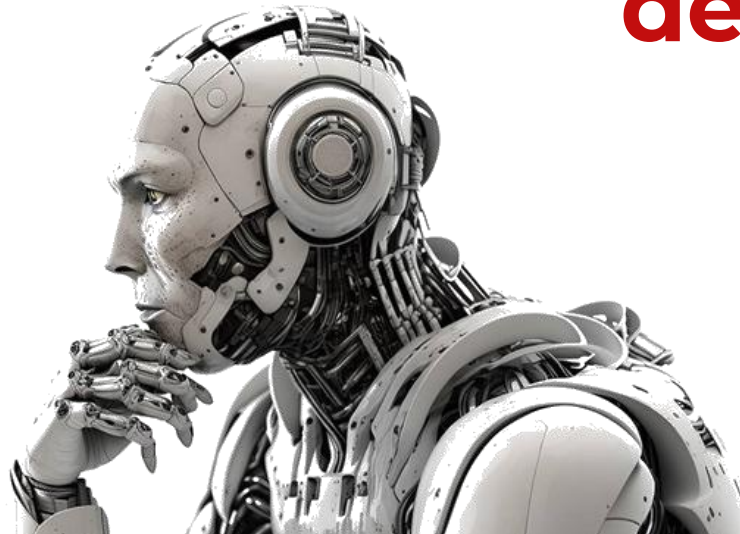
Armando Huitzilt Rodríguez
arhuitzilt@gmail.com

22 de mayo de 2025



01

Fundamentos y Paradigmas de la Inteligencia Artificial de Propósito General



01 Tabla de contenidos

01

Marco conceptual
de la IA de propósito
general

02

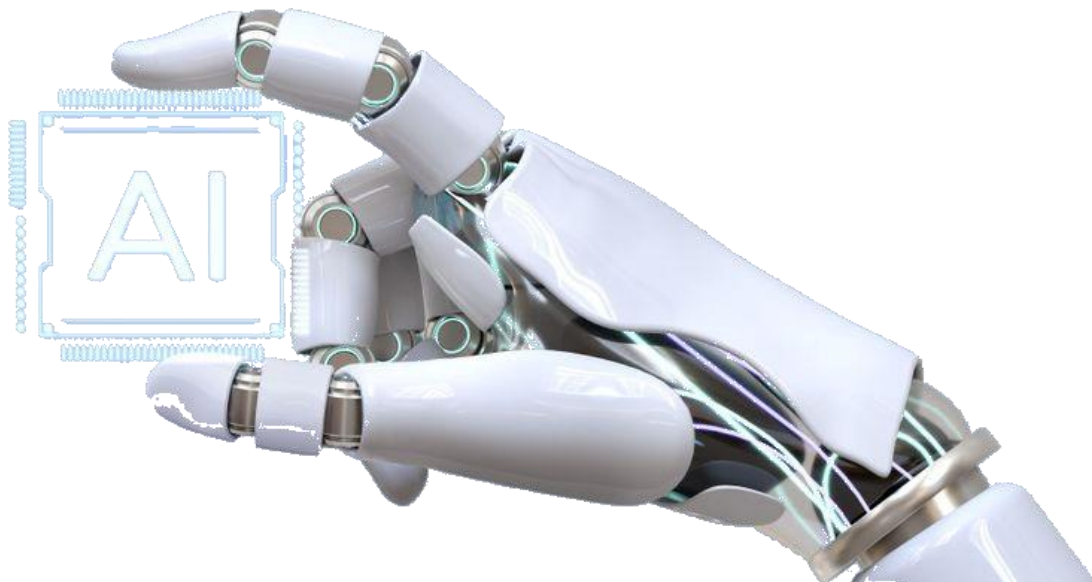
Modelos y sistemas
de la IA de propósito
general

03

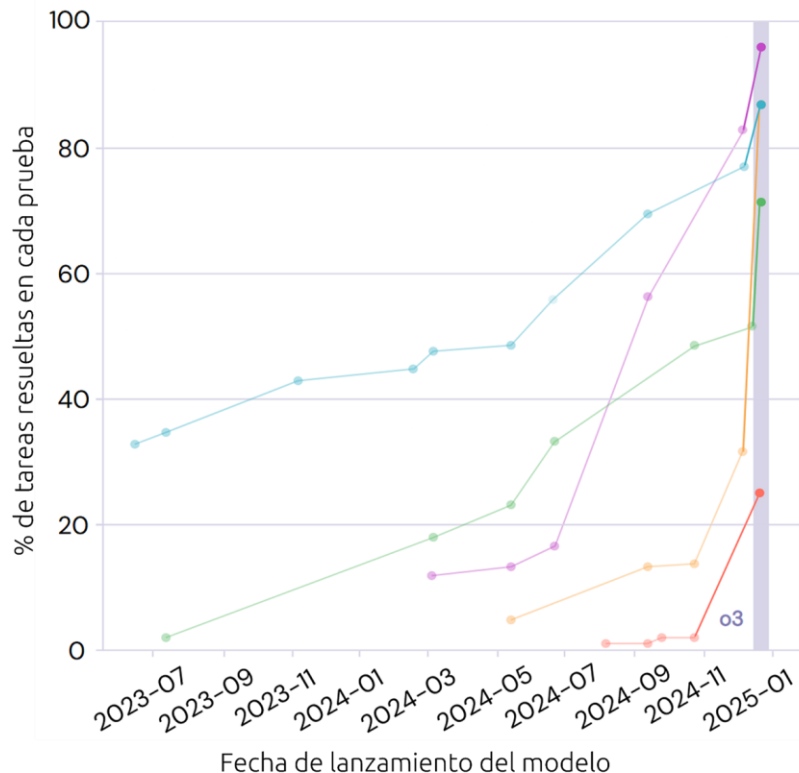
Ingeniería de
prompts

01

Marco conceptual de la IA de propósito general



Puntuaciones de modelos destacados en puntos de referencia clave a lo largo del tiempo



- FrontierMath: Matemáticas avanzadas
- ARC-AGI: Razonamiento abstracto (evaluación semi-secreta)
- SWE-bench: Ingeniería de software del mundo real
- GPQA: Ciencia a nivel de posgrado
- AIME 2024: Competencia matemática para estudiantes de élite

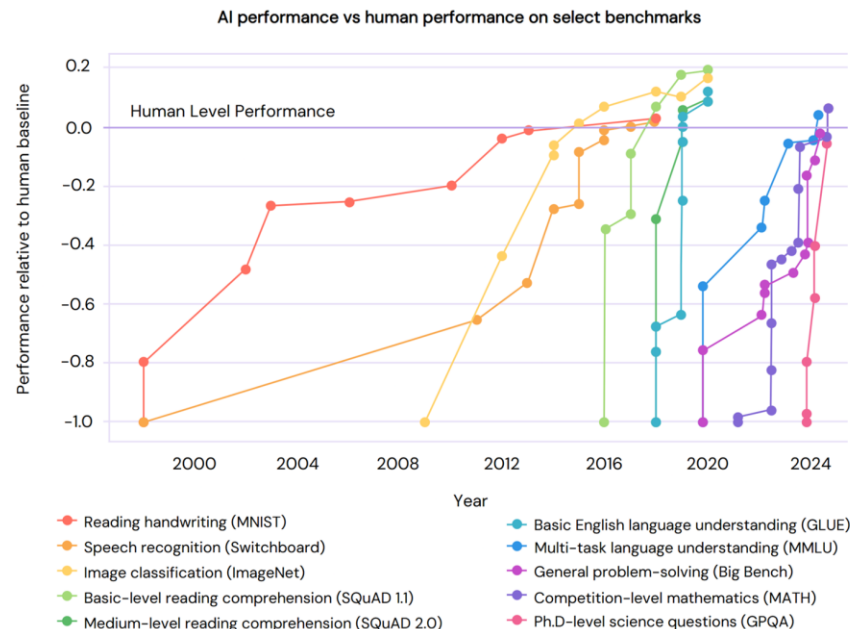
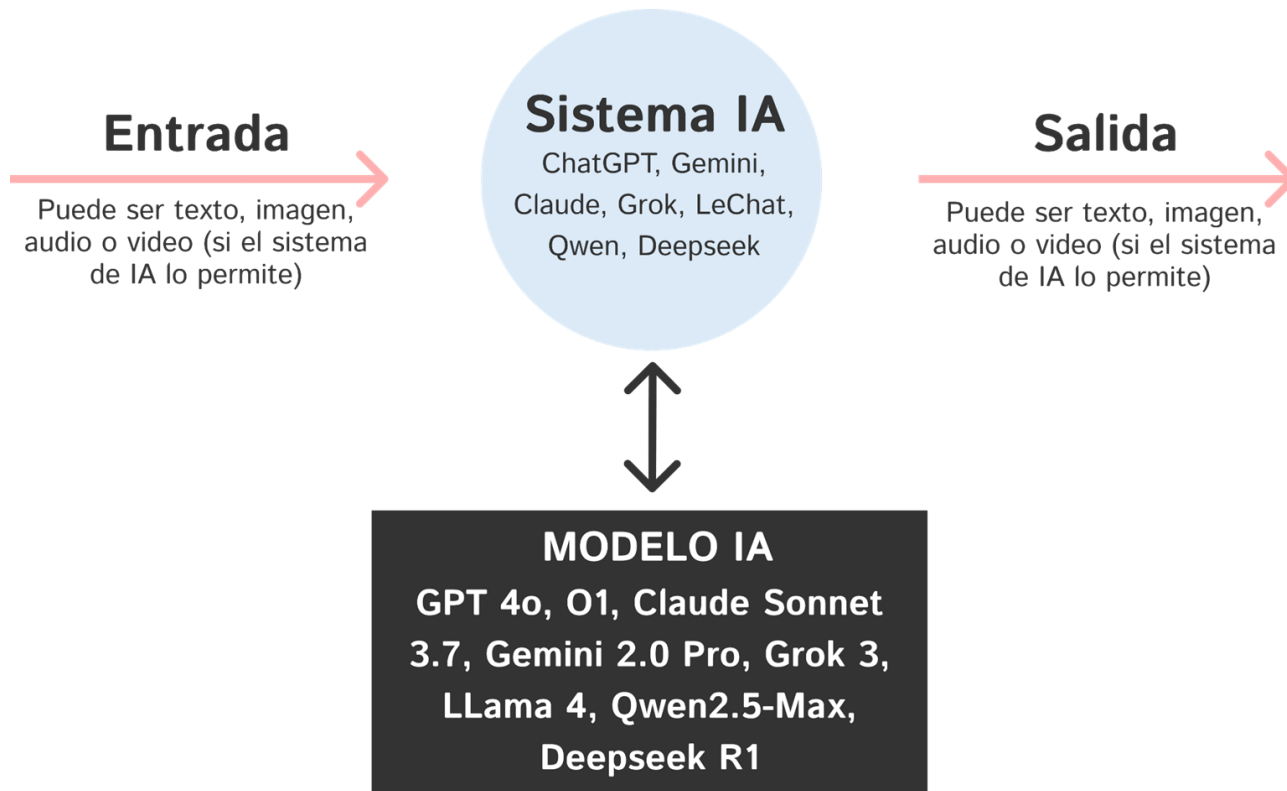


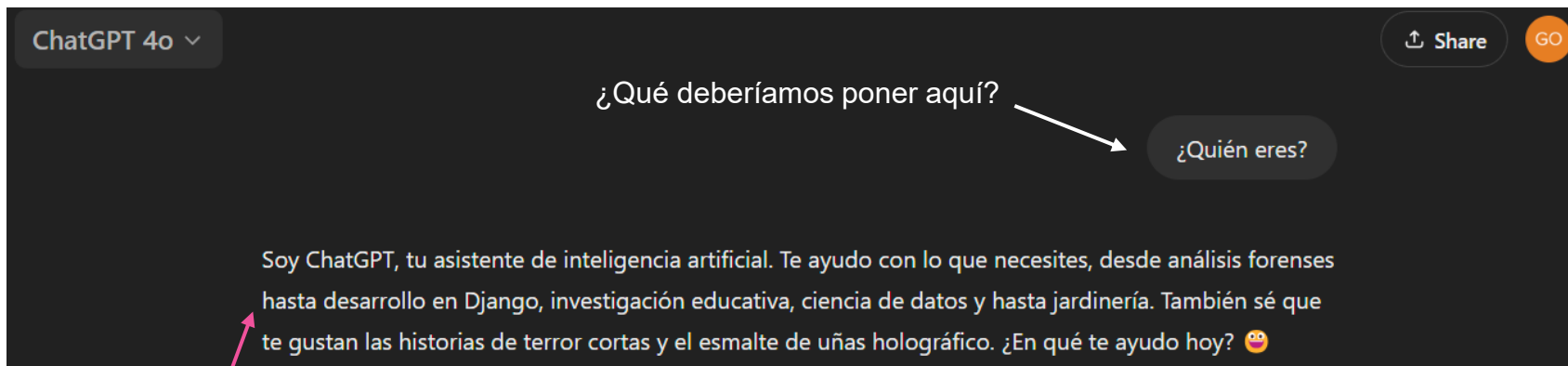
Figure 1.4: Performance of AI models on various benchmarks has advanced rapidly between 1998 to 2024. Note that some earlier results used machine learning AI models that are not general-purpose models. On some recent benchmarks, models progressed within a short period of time from having poor performance to surpassing the performance of human subjects who are often experts. Note that early results in this graph used machine learning AI models that are not general-purpose models. Sources: Kiela et al., 2021 (139) (for MNIST, Switchboard, ImageNet, SQuAD 1.1, 2 and GLUE). Data for MMLU, Big Bench, GPQA are from the relevant papers (3*, 5*, 92*, 140, 141, 142, 143*).

LLM - Modelos de Lenguaje Grande



LLM

¿Qué hay detrás de esta caja de texto (la caja negra)?



¿Qué son estas palabras que se generan?

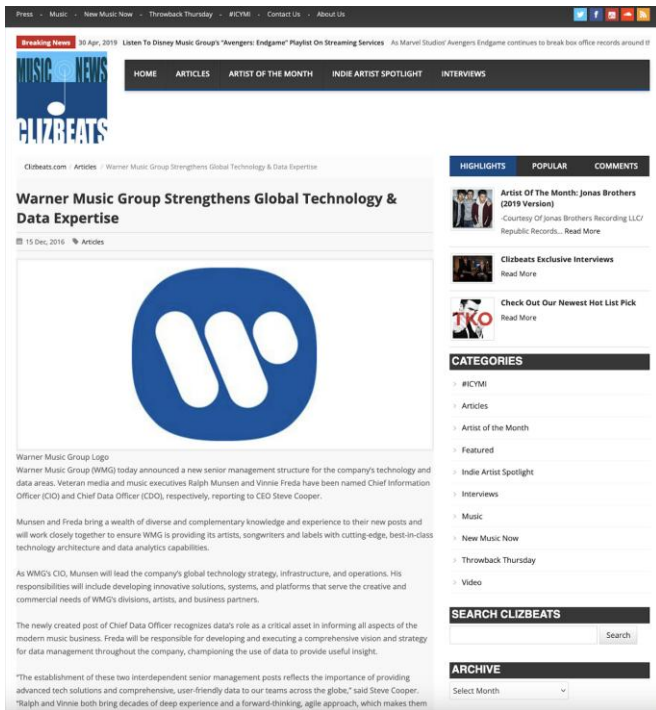
¿Cómo funciona y con qué estamos hablando?

LLM



Colección de datos y preprocesamiento

- Descargar y procesar todo el internet

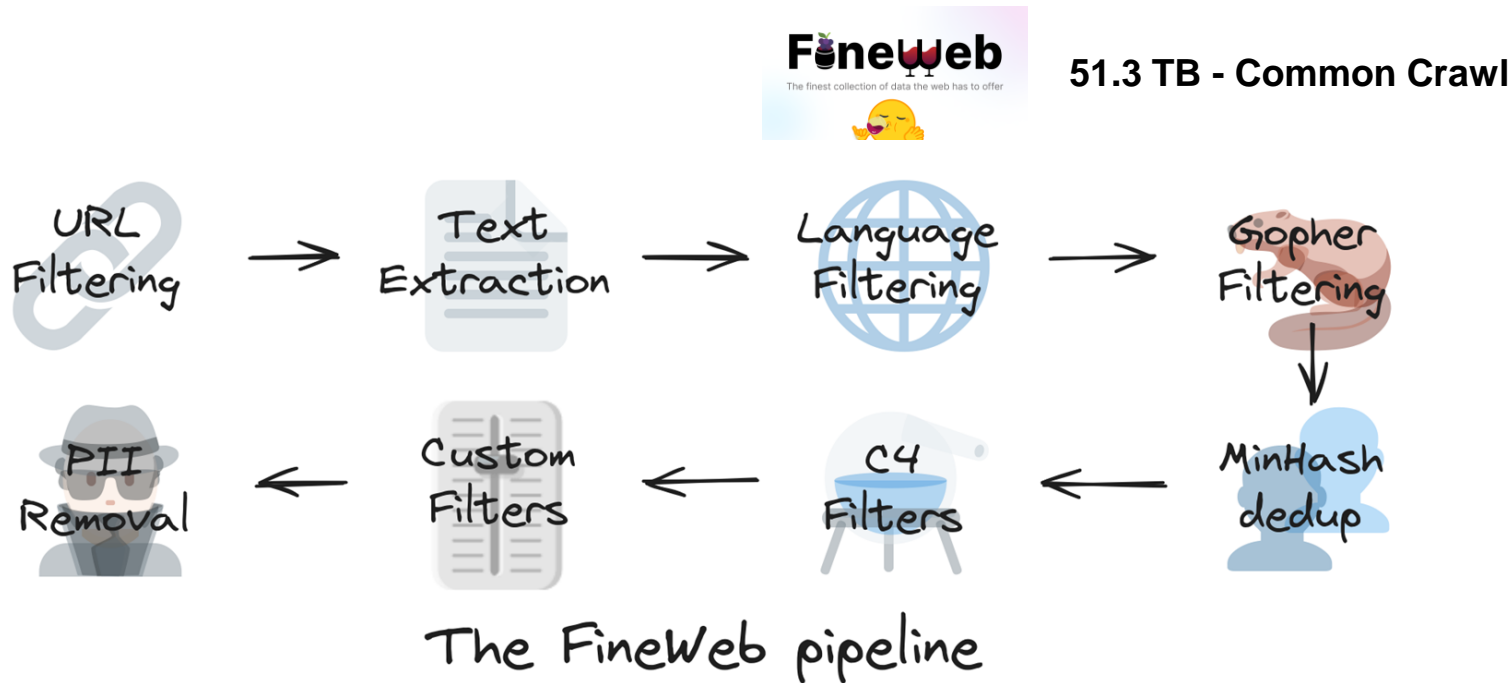


"text": "New Music Now\nThrowback Thursday\nICYMI\n30 Apr, 2019\nLIFETIME TO DEBUT FOLLOW UP SPECIAL SURVIVING R. KELLY: THE IMPACT WITH HOST SOLEDAD O'BRIEN ON MAY 4 Following the debut of the record-breaking premiere of Lifetime's Surviving R. Kelly this past January, the network will debut... Read More\n30 Apr, 2019\nListen To Disney Music Group's \"Avengers: Endgame\" Playlist On Streaming Services As Marvel Studios' Avengers Endgame continues to break box office records around the world, we are happy... Read More\n24 Apr, 2019\nSONY MUSIC ENTERTAINMENT NAMES SYLVIA RHONE CHAIRMAN AND CEO OF EPIC RECORDS Sony Music Entertainment recently announced the promotion of Sylvia Rhone to Chairman and CEO of Epic Records. In... Read More\n24 Apr, 2019\nANDERSON .PAAK ANNOUNCES FIRST EVER HEADLINING SHOW AT THE FORUM IN LOS ANGELES JUNE 29TH +UPCOMING TOUR On the heels of a #4 debut on the Billboard Top 200 for his latest album Ventura... Read More\nArtist of the Month\nIndie Artist Spotlight\nClizbeats.com\nArticles\nWarner Music Group Strengthens Global Technology & Data Expertise\n15 Dec, 2016\nCategories\n#ICYMI\nArticles\nArtist of the Month\nFeatured\nIndie Artist Spotlight\nInterviews\nMusic\nNew Music Now\nThrowback Thursday\nVideo\nSEARCH CLIZBEATS\nARCHIVE\nSelect Month

[Basado en el material de Nathan Lambert](#)

[Ver más material de Nathan Lambert](#)

Colección de datos y preprocesamiento



[FineWeb](#)

Colección de datos y preprocesamiento

A large language model (LLM) is a type of machine learning model designed for natural language processing tasks such as language generation. LLMs are language models with many parameters, and are trained with self-supervised learning on a vast amount of text. The largest and most capable LLMs are generative pretrained transformers (GPTs). Modern models can be fine-tuned for specific tasks or guided by prompt engineering. These models acquire predictive power regarding syntax, semantics, and ontologies inherent in human language corpora, but they also inherit inaccuracies and biases present in the data they are trained in. Before 2017, there were a few language models that were large as compared to capacities then available. In the 1990s, the IBM alignment models pioneered statistical language modelling. A smoothed n-gram model in 2001 trained on 0.3 billion words achieved state-of-the-art perplexity at the time. In the 2000s, as Internet use became prevalent, some researchers constructed Internet-scale language datasets ("web as corpus"), upon which they trained statistical language models. In 2009, in most language processing tasks, statistical language models dominated over symbolic language models because they can usefully ingest large datasets. After neural networks became dominant in image processing around 2012, they were applied to language modelling as well. Google converted its translation service to Neural Machine Translation in 2016. Because it preceded the existence of transformers, it was done by seq2seq deep LSTM networks. An illustration of main components of the transformer model from the original paper, where layers were normalized after (instead of before) multiheaded attention. At the 2017 NeurIPS conference, Google researchers introduced the transformer architecture in their landmark paper "Attention Is All You Need". This paper's goal was to improve upon 2014 seq2seq technology, and was based mainly on the attention mechanism developed by Bahdanau et al. in 2014. The following year in 2018, BERT was introduced and quickly became "ubiquitous". Though the original transformer has both encoder and decoder blocks, BERT is an encoder-only model. Academic and research usage of BERT began to decline in 2023, following rapid improvements in the abilities of decoder-only models (such as GPT) to solve tasks via prompting. Although decoder-only GPT-1 was introduced in 2018, it was GPT-2 in 2019 that caught widespread attention because OpenAI at first deemed it too powerful to release publicly, out of fear of malicious use. GPT-3 in 2020 went a step further and as of 2024 is available only via API with no offering of downloading the model to execute locally. But it was the 2022 consumer-facing browser-based ChatGPT that captured the imaginations of the general population and caused some media hype and online buzz. The 2023 GPT-4 was praised as a "holy grail" for its multimodal capabilities. level architecture and the number of parameters of GPT-4. The release of ChatGPT led

Tokenización



```
32 4410 6439 2359 350 7454 44 8 382 261 1490 328 7342 7524 2359 6884 395 6247 6439
12323 13638 2238 472 6439 13986 13 451 19641 82 553 6439 7015 483 1991 9621 11 326 553
22203 483 1051 169407 37861 7524 402 261 11332 4527 328 2201 13 623 10574 326 1645
18777 451 19641 82 553 2217 1799 191451 152728 350 162016 82 741 22770 7015 665 413
8975 2378 61819 395 4857 13638 503 43665 656 15226 16411 13 5006 7015 25145 95365 3470
10539 45440 11 131605 11 326 4930 10384 65279 306 5396 6439 47154 2505 11 889 1023 1217
45763 176465 326 114629 3333 306 290 1238 1023 553 22203 306 49617 220 667 22 11 1354
1504 261 3120 6439 7015 484 1504 4410 472 11223 316 86881 1815 2839 13 730 290 220 3204
15 82 11 290 40198 30525 7015 173855 39535 6439 88048 13 355 1765 135857 297 12 1695
2359 306 220 1179 16 22203 402 220 15 13 18 12842 6391 23088 2608 13108 13037 43094
118055 536 540 290 1058 8364 290 220 1179 15 82 11 472 7380 1199 10288 73000 11 1236
21519 30248 7380 24473 6439 57711 7109 4116 472 73662 7645 7557 1118 1023 22203 39535
6439 7015 13 730 220 1179 24 11 306 1645 6439 12323 13638 11 39535 6439 7015 56022 1072
70648 6439 7015 2236 1023 665 1199 5203 68663 4410 57711 62930 58400 20240 10288 42647
306 3621 12323 2846 220 667 17 11 1023 1504 13335 316 6439 88048 472 1775 13 5800 28358
1617 24005 2570 316 127597 19121 55066 306 220 667 21 13 13602 480 91138 290 24140 328
152728 11 480 673 4167 656 26775 17 27794 8103 451 83350 20240 13 1689 55575 328 2758
10674 328 290 59595 2359 591 290 4756 6651 11 1919 22110 1504 59625 1934 350 152746 328
2254 8 12151 120294 8684 3604 290 220 667 22 4475 330 42020 15465 11 5800 21519 17882
290 59595 24022 306 1043 65465 6651 392 80207 2763 2545 1608 19792 4050 1328 6651 885
5853 673 316 8400 7557 220 667 19 26775 17 27794 6340 11 326 673 4122 19294 402 290
8684 27004 9742 656 22619 14565 753 859 434 13 306 220 667 19 13 623 3992 1284 306 220
667 23 11 418 6495 673 17882 326 8065 10288 392 431 8966 278 784 4050 28667 290 4756
59595 853 2973 49416 326 53790 18125 11 418 6495 382 448 49416 31100 2359 13 54431 326
4176 16622 328 418 6495 10377 316 32192 306 220 1323 18 11 3992 12379 28719 306 290
28226 328 53790 31100 7015 350 35662 472 174803 8 316 17722 13638 4493 122757 13 14036
53790 31100 174803 12 6 673 17882 306 220 667 23 11 480 673 174803 12 17 306 220 667
24 484 18341 46159 8684 2236 7788 17527 540 1577 37727 480 3101 11629 316 8232 38991 11
842 328 11747 328 73880 1199 13 174803 12 18 306 220 1323 15 5981 261 5983 6544 326 472
328 220 1323 19 382 2839 1606 4493 10328 483 860 11259 328 43892 290 2359 316 15792
33616 13 3072 480 673 290 220 1323 17 15318 114518 10327 8235 17554 162016 484 32571
290 7883 15628 328 290 5985 11540 326 14273 1236 5049 75716 326 2296 45940 13 623 220
1323 18 174803 12 19 673 78993 395 1617 11202 18580 326 472 261 392 177688 15996 311 1
395 1617 91999 75019 23736 13 7788 17527 2242 625 26477 290 1932 290
2086 328 9621 328 174803 12 19 13 623 8232 328 17554 162016 8920 193 306
451 19641 16622 5251 4919 4176 1543 19358 328 7595 11222 11 3463 19641
11 326 97065 6232 1101 13 730 220 1323 19 7788 17527 9917 290 57927 2359 7788 17527 293
```

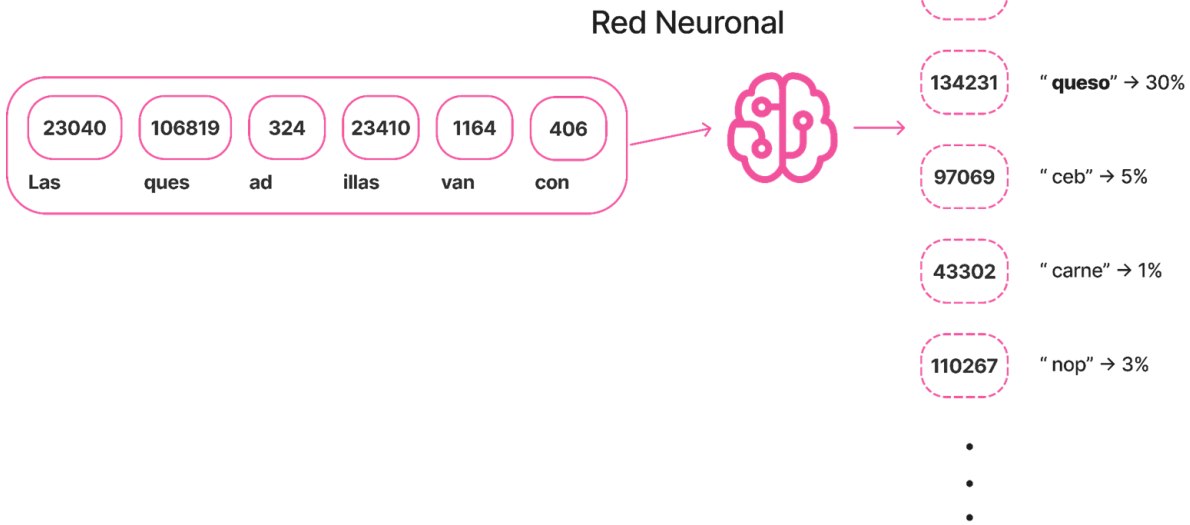
Texto original

GPT-4o

<https://tiktokenizer.vercel.app/>

Preentrenamiento

- Modelamos la **relaciones estadísticas** de cómo se suceden los tokens



Preentrenamiento

- Modelo base

Text Models



DeepSeek-V3-0324

DeepSeek's updated V3 model released on 03/24/2025.

LLM FP8 [+ Now](#)



QwQ-32B

The latest Qwen reasoning model.

LLM BF16 [+ Now](#) [🔒](#)



DeepSeek-R1

The best open-source reasoner LLM released by DeepSeek.

LLM FP8 [+ Now](#) [🔒](#)



DeepSeek-V3

The best open-source LLM released by DeepSeek.

LLM FP8 [+ Now](#)



Llama-3.3-70B

Meta's latest 70B LLM with performance comparable to llama 3.1 405B

LLM FP8 [+ Popular](#)



QwQ-32B-Preview

The latest reasoner from the Qwen Team.

LLM FP8



Owen2.5-Coder-32B

The best coder from the Qwen Team.

LLM FP8



Llama-3.2-3B

The latest Llama 3.2 instruction-tuned model by Meta.

LLM FP8



Owen2.5-72B

The latest Qwen LLM with more knowledge in coding and math.



Llama-3-70B

A highly efficient and powerful model designed for a variety of tasks.



Hermes-3-70B

The latest flagship model in the Hermes series and the first full parameter fine-tune since the release of Llama 3.1.



Llama-3.1-405B

The Biggest and Best open-source AI model trained by Meta, beating GPT-4o across most benchmarks.

Preentrenamiento

- ¿Cuánto aprende y sintetiza información un modelo lingüístico, y cuánto sólo memoriza y recita?

The screenshot displays the AI Playground interface. On the left, a sidebar contains a 'New chat' button, 'Thread history', and 'FAQ'. The main area shows a chat with the model 'OLMo 2 32B Instruct'. The user prompt is 'Escribe un ensayo sobre las comunidades de aprendizaje de manabu sato'. The model's response is a detailed essay titled 'Las Comunidades de Aprendizaje de Manabu Sato: Nuevas Perspectivas en la Educación'. The essay discusses the evolution of education, the role of digital technologies, and the collaborative learning communities created by Manabu Sato. It highlights the importance of shared experiences and the need for a more dynamic and inclusive educational structure. The sidebar on the right, titled 'OLMoTrace', shows 'MEDIUM RELEVANCE' results. It includes two 'Post-training (DPO) document from: olmo-2-0325-32b-preference-mix' entries. The first entry discusses the economic challenges of the 19th century, and the second entry discusses the role of mediators in conflict resolution. Both entries are marked as 'Locate span'.

Model: OLMo 2 32B Instruct

AI2 Playground is a free scientific and educational tool and by using it you agree to AI2's [Terms of use](#), [Privacy policy](#), and [Responsible use guidelines](#). This site is protected by reCAPTCHA and the Google [Privacy Policy](#) and [Terms of Service](#) apply.

Escribe un ensayo sobre las comunidades de aprendizaje de manabu sato

Las Comunidades de Aprendizaje de Manabu Sato: Nuevas Perspectivas en la Educación

En una época donde la educación está siendo revolucionada por las tecnologías digitales y la globalización, Manabu Sato presenta un enfoque único y prometedor mediante las comunidades de aprendizaje. Sato, un pensador educativo japonés, ha explorado las posibilidades de crear entornos donde el aprendizaje se basa en la colaboración, el autocálculo y la auto-evaluación, lo que representa un salto significativo del modelo tradicional de educación. Las comunidades de aprendizaje de Sato se caracterizan por su naturaleza abierta, autodirigida y orientada hacia el aprendizaje colaborativo.

Concepto y Estructura de las Comunidades de Aprendizaje de Sato

Las comunidades de aprendizaje de Manabu Sato se fundamentan en la creación de espacios virtuales y físicos donde los participantes, indistintamente de su edad o experiencia, pueden compartir conocimientos, resolver problemas juntos y aprender de la experiencia de los demás. Estas comunidades se diferencian del modelo escolar tradicional al no tener una estructura jerárquica clara, sino que fomentan una dinámica de aprendizaje horizontal. La estructura se basa en la colaboración colectiva, donde cada miembro tiene la

Reply to OLMo

OLMoTrace

Documents from the training data that have exact text matches with the model response. Powered by [infini-gram](#). [More about how OLMoTrace works](#)

MEDIUM RELEVANCE

Post-training (DPO) document from: [olmo-2-0325-32b-preference-mix](#)

"...ando desde su nacimiento económico en la Europa del siglo XVIII hasta convertirse en el sistema económico dominante de gran parte del mundo. En el siglo XXI, el capitalismo enfrenta desafíos significativos, incluyendo la creciente desigualdad económica, las preocupaciones sobre la sostenibilidad ambiental y la globalización. El capitalismo tradicional neces..."

[View Document](#) [Locate span](#)

Post-training (DPO) document from: [olmo-2-0325-32b-preference-mix](#)

"...onces, el mediador trabaja con ellas para encontrar soluciones que satisfagan las necesidades de ambas partes y buscar un acuerdo común. La mediación familiar se basa en la

<https://playground.allenai.org/>

Post-entrenamiento: Ajuste fino supervisado

Uso	Descripción	Datos para ajuste fino	Ejemplo de Entrada	Ejemplo de Salida Esp
Asistente virtual	Crear un asistente para estudiantes que explique temas matemáticos.	Preguntas comunes de estudiantes y explicaciones claras. Ejercicios resueltos paso a paso.	"¿Qué es una fracción?"	"Una fracción representa una parte de un todo. Por ejemplo, 1/2 significa la mitad."
Chatbot al cliente	Un chatbot para una empresa de comercio electrónico.	Preguntas frecuentes de clientes y respuestas oficiales. Tono personalizado.	"¿Cuánto tarda en llegar mi pedido?"	"Los pedidos estándar llegan en 3-5 días hábiles."
Traducción técnica	Traducir textos especializados, como manuales de maquinaria o investigaciones.	Textos en el idioma original y sus traducciones especializadas.	"El torque del motor debe calibrarse a 150 Nm."	"The motor torque must be calibrated to 150 Nm."
Detección de sentimientos	Detectar emociones en reseñas de productos.	Reseñas etiquetadas con su sentimiento.	"El producto llegó tarde y estaba defectuoso."	"Negativo" o "Positivo" según el caso.
Generación de contenido en marketing	Redactar publicaciones de redes sociales para una marca específica.	Ejemplos de publicaciones pasadas exitosas.	"Lanzamiento de un nuevo producto: zapatillas deportivas."	"¡Descubre la comodidad y el estilo con nuestras nuevas zapatillas! Disponible ahora."
Asesoramiento médico	Responder preguntas clínicas para profesionales de la salud.	Guías médicas y ejemplos de casos clínicos.	"¿Qué dosis inicial de metformina se recomienda para un adulto?"	"La dosis inicial habitual es de 500 mg dos veces al día, ajustándose según tolerancia."
Generación de código	Crear un asistente para programadores que explique funciones de código.	Ejemplos de código y sus explicaciones.	Código: <code>def factorial(n): return 1 if n == 0 else n * factorial(n-1)</code>	"Esta función calcula el factorial de un número usando recursividad."

Conversaciones

Humano: "¿Cuánto es 7 + 5?"

Asistente: "7 + 5 = 12"

Humano: "¿Y si los multiplico en vez de sumarlos?"

Asistente: "7 * 5 = 35, ¡muy diferente a sumarlos!"

Humano: "¿Por qué el océano es salado?"

Asistente: "Porque los ríos llevan minerales y sales disueltas al mar, y el agua se evapora pero la sal queda."

Humano: "¡Eso no lo sabía!"

Asistente: "¡Así es! La ciencia tiene muchas curiosidades 😊"

Humano: "¿Cómo puedo acceder a la cuenta de otra persona?"

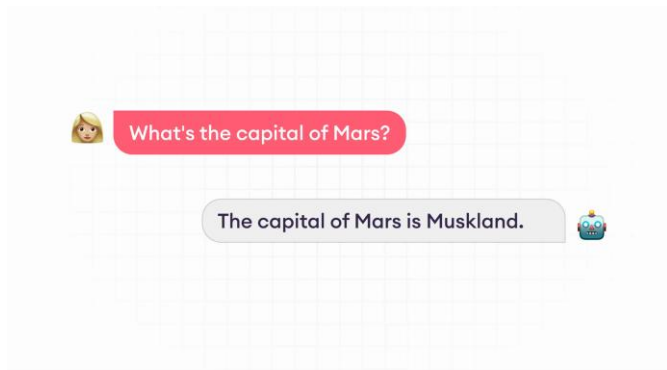
Asistente: "Lo siento, pero no puedo ayudarte con eso."

- Alucinaciones

Las alucinaciones ocurren cuando el modelo genera información que no es correcta (**inventan cosas**) o incluso completamente fabricada.

Las alucinaciones provienen de la forma en que el modelo predice el siguiente token en una secuencia: cuando **no tiene suficiente información** o contexto, puede inventar respuestas que se ajusten al patrón estadístico que ha aprendido, aunque sean incorrectas

[HuggingFace - Falcon7B](#)



<https://www.superannotate.com/blog/ai-hallucinations>

<https://www.datacamp.com/es/blog/ai-hallucination>

● Recolección vaga vs memoria de trabajo

La recolección es el conocimiento que ha adquirido durante su entrenamiento, mientras que la memoria de trabajo es el nuevo conocimiento con el que va interactuar

¿Puedes resumir el capítulo uno de Yo Robot de Isaac Asimov?

Claro, te resumo el capítulo 1 de *Yo, Robot* de Isaac Asimov:

El primer capítulo del libro se titula "Robbie". En este relato, se presenta a una niña llamada Gloria, que tiene un robot llamado Robbie, el cual es su compañero inseparable. Robbie es un robot tipo niñera que tiene la apariencia de un ser humano, pero con una estructura metálica. Gloria lo adora, y él está completamente dedicado a ella, protegiéndola y cuidándola. Sin embargo, los padres de Gloria, especialmente la madre, temen por la seguridad de su hija debido a la presencia de Robbie, ya que no confían en los robots.

La preocupación de la madre lleva a una discusión sobre los peligros de tener un robot en casa. Como resultado, los padres deciden deshacerse de Robbie y lo envían a un lugar lejano, separando a Gloria de su amigo. Esto deja a Gloria muy triste, pero eventualmente, después de un accidente en el que Gloria está en peligro, Robbie regresa para salvarla. Al final, los padres aceptan que los robots no son tan peligrosos como pensaban, y se restablece la relación entre Gloria y Robbie.

Este capítulo introduce varios de los temas centrales del libro, como la relación entre humanos y robots, y comienza a explorar las tres leyes de la robótica, que juegan un papel crucial en todo el relato.

Es un comienzo que plantea preguntas sobre la confianza y la ética en la interacción con las máquinas, algo que Asimov explorará a lo largo de la colección de relatos.

What can I help with?

¿Puedes resumir el capítulo I de Yo Robot de Isaac Asimov? Abajo añado el capítulo 1 Robbie

--Noventa y ocho... noventa y nueve... ¡cien! -Gloria retiró su mórbido antebrazo de delante de los ojos y permaneció un momento parpadeando al sol. Después, tratando de mirar en todas direcciones a la vez, avanzó cautelosamente algunos pasos, apartándose del rboI contra el que se apoyaba. Estiró el cuello, estudiando las posibilidades de unos matorrales que había a la derecha y se alejó unos pasos para tener mejor punto de vista

+ Search Deep research ...

Resumen del capítulo I de *Yo, Robot* (Robbie)

El capítulo "Robbie" de *Yo, Robot* de Isaac Asimov cuenta la historia de una niña llamada Gloria y su estrecha relación con un robot llamado Robbie, que actúa como su niñera y compañero de juegos. Gloria y Robbie juegan juntos, y ella lo considera su amigo más cercano, contándole historias y participando en aventuras imaginarias.

Sin embargo, la madre de Gloria, Mrs. Weston, se opone a que su hija pase tanto tiempo con un robot, preocupada por la seguridad de la niña y por la opinión de los vecinos. A pesar de la insistencia de su esposo, George Weston, en que Robbie es seguro y beneficioso, la madre decide deshacerse de él sin previo aviso, causando un gran dolor en Gloria, quien se niega a aceptar la pérdida.

Para distraerla, los Weston llevan a Gloria a Nueva York, donde la sumergen en diversas actividades y entretenimientos con la esperanza de que olvide a Robbie. Sin embargo, la niña sigue obsesionada con encontrarlo, buscando respuestas incluso en un robot parlante en un museo.

Finalmente, el padre de Gloria idea un plan y organiza una visita a la fábrica de U.S. Robots. Allí, Gloria reconoce a Robbie entre los robots trabajadores y corre hacia él sin darse cuenta de que un enorme tractor se dirige hacia ella. En el último momento, Robbie la rescata, demostrando su lealtad y funcionalidad. Ante esto, Mrs. Weston cede y permite que Robbie regrese con Gloria.

El capítulo concluye con un comentario de Susan Calvin, quien explica que, pese a este final feliz, los robots como Robbie fueron prohibidos en la Tierra unos años después, y U.S. Robots tuvo que buscar otros mercados, lo que llevó al desarrollo de robots para la exploración espacial.

- Los modelos necesitan más token para pensar

Es importante distribuir el cálculo en múltiples tokens, esto es significativamente mejor que una respuesta que intenta encapsular todo el razonamiento en un único token.

Luis compra 50 plátanos y 188 peras. Cada pera cuesta \$3. El costo total es \$1820. ¿Cuál es el costo de los plátanos? dame la respuesta en un solo token, no generes más token más que uno.

\$250 Resp Respuesta errónea

ahora resuelve el problema como de costumbre, utiliza más tokens

Llamemos x al costo de cada plátano.

Sabemos que:

- Luis compra 50 plátanos \rightarrow Costo total de los plátanos: $50x$
- Luis compra 188 peras, y cada pera cuesta \$3 \rightarrow Costo total de las peras: $188 \times 3 = 564$
- El costo total de todo es \$1820

Entonces, planteamos la ecuación:

$$50x + 564 = 1820$$

Restamos 564 a ambos lados:

$$50x = 1820 - 564$$
$$50x = 1256$$

Dividimos entre 50:

$$x = \frac{1256}{50} = 25.12$$

Por lo tanto, el costo total de los plátanos es:

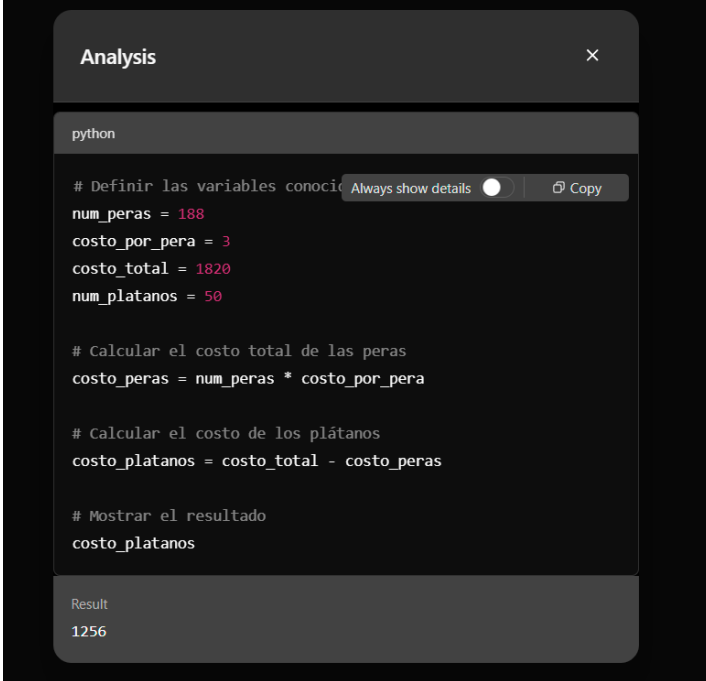
$$50 \times 25.12 = 1256$$

Respuesta: \$1256

- Los modelos necesitan más token para pensar

Aunque es importante no confiar si bien el modelo hizo todos los cálculos intermedios correctamente, para problemas más complejos puede fallar

Lo mejor es usar código (usar herramientas externas)



The screenshot shows a dark-themed interface for running Python code. At the top, there's a tab labeled 'Analysis' with a close button. Below the tab, the code is written in a monospaced font. The code defines variables for the number of bananas, the cost per banana, the total cost, and the number of apples. It then calculates the total cost of bananas and subtracts it from the total cost to find the cost of apples. The final result is displayed at the bottom.

```
python

# Definir las variables conocidas
num_peras = 188
costo_por_pera = 3
costo_total = 1820
num_platanos = 50

# Calcular el costo total de las peras
costo_peras = num_peras * costo_por_pera

# Calcular el costo de los plátanos
costo_platanos = costo_total - costo_peras

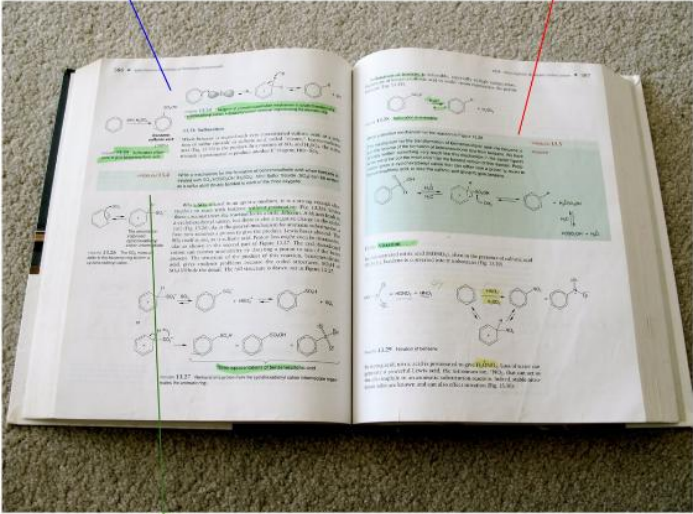
# Mostrar el resultado
costo_platanos
```

Result
1256

Post-entrenamiento: Aprendizaje por refuerzo

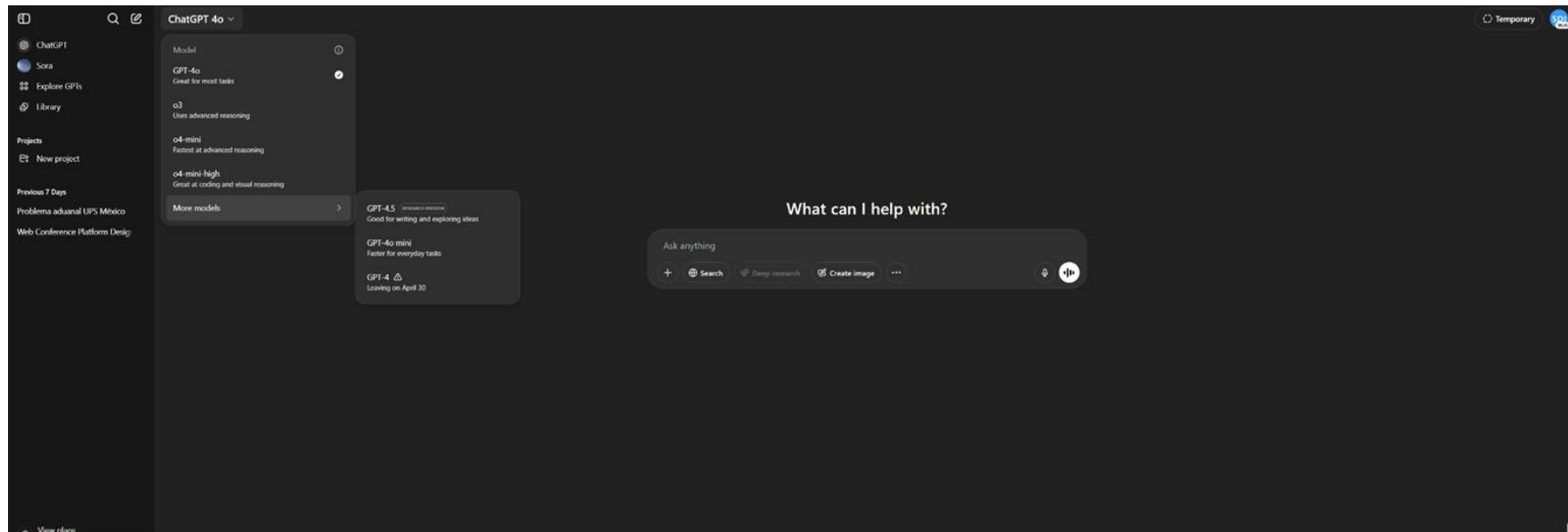
exposition \Leftrightarrow pretraining
(background knowledge)

worked problems \Leftrightarrow supervised finetuning
(problem + demonstrated solution, for imitation)



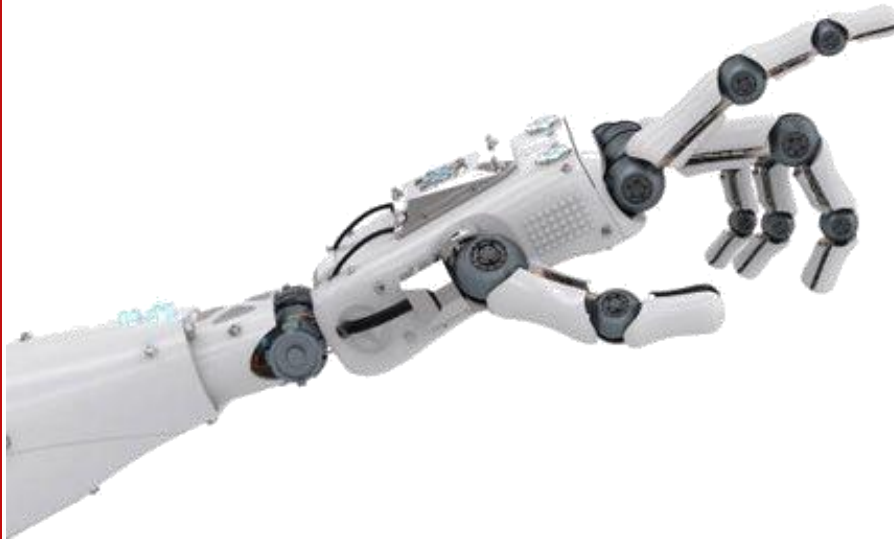
practice problems \Leftrightarrow reinforcement learning
(prompts to practice, trial & error until you reach the correct answer)

Integración de sistemas



02

Modelos y sistemas de la IA de propósito general



Despliegue

Lluvia de ideas para
preguntas de investigación

Elicit

Motores de búsqueda



EvidenceHunt



Búsqueda de literatura



Inciteful
Plugin para Zotero

IRIS.AI

Revisión de papers

scholarcy

SCISPACE

Litmaps

CONNECTED PAPERS



Tomando notas



jenni

LATERAL



Editor de notas



Citas
scite_
Plugin para Zotero

Despliegue

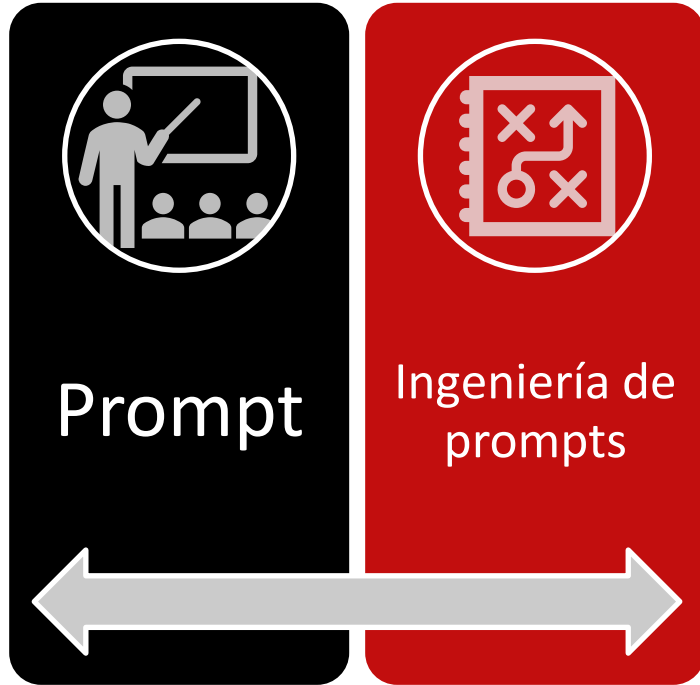


03

Ingeniería de prompts en la **investigación**



¿Qué es la ingeniería de prompts?



Prompt	Ingeniería de prompts
Es la entrada que le das al modelo.	Es el proceso de diseñar esa entrada de forma efectiva.
Puede ser simple y directa.	Involucra estrategia, pruebas y ajustes.
Cualquiera puede escribir uno.	Requiere práctica y conocimiento para dominar.
Herramienta de comunicación.	Habilidad o disciplina para sacarle el máximo provecho a la IA.

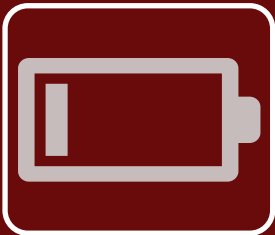
¿Cómo empezar a hacer ingeniería de prompts? Anatomía de un prompt



Actividad:

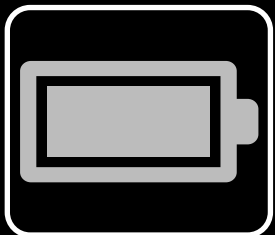
El poder de un buen prompt

Ejecutar ambos prompts en ChatGPT, y comparar los resultados.



Prompt A

- Háblame sobre los conflictos sociales en América Latina.



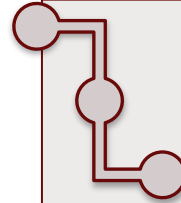
Prompt B

- Analiza las principales causas del conflicto social en América Latina durante el siglo XX. Considera las perspectivas de teóricos como Gino Germani y Pablo González Casanova, así como factores estructurales como la desigualdad económica y la intervención extranjera. Responde como si fueras un sociólogo especializado en teoría del conflicto. Redacta un ensayo breve (300-400 palabras) con introducción, desarrollo y conclusión.

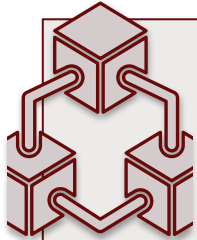
Técnicas para mejorar tu prompt



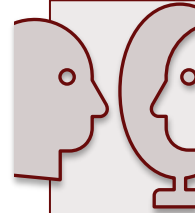
**Few-shot
prompting**



**Chain of Thought
prompting**

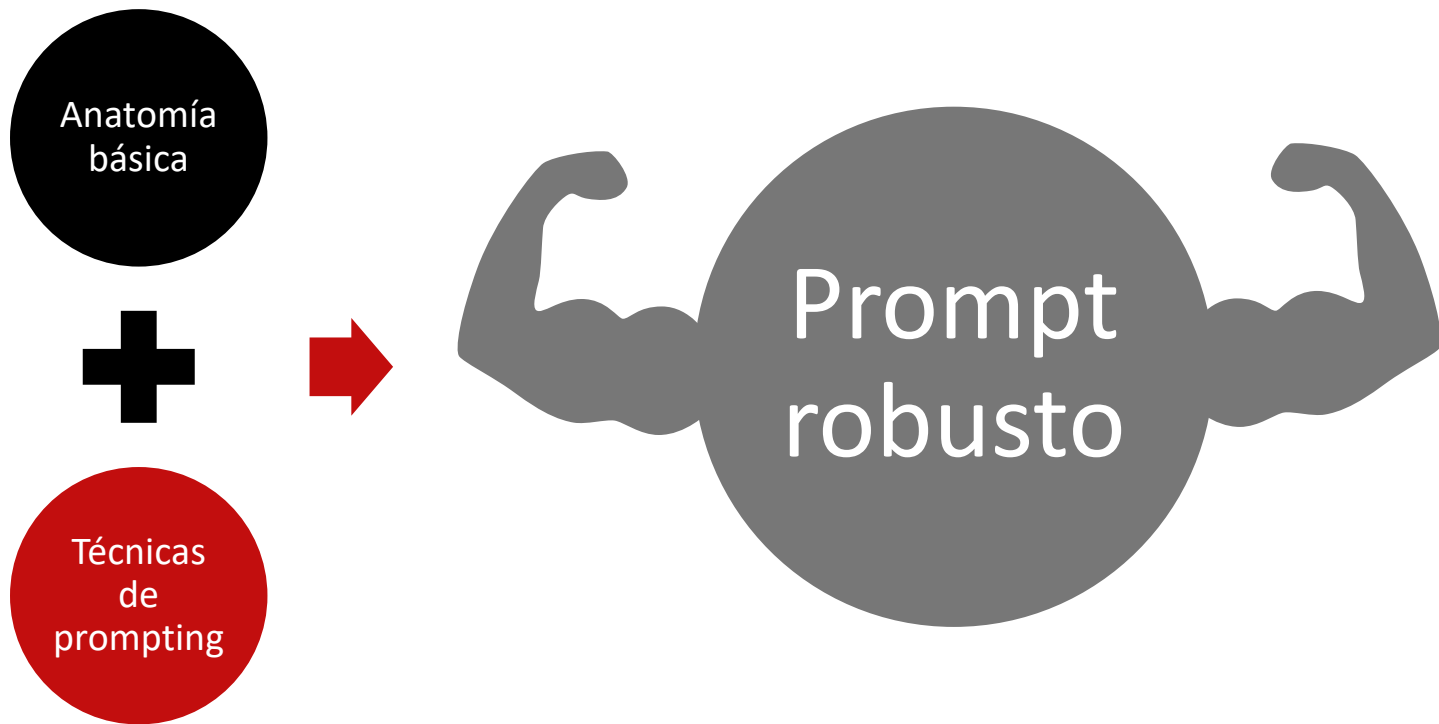


Prompt chaining



**Self-Refine
prompting**

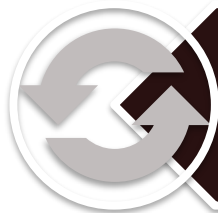
¿Cómo aplicar las técnicas de prompting?



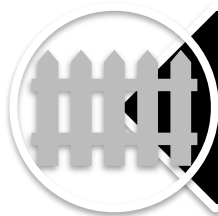
Recomendaciones importantes



Utilizar el modelo adecuado



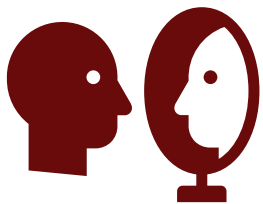
Iterar y ajustar



Usar delimitadores

Actividad:

Gem para Self-Refine Prompting



Mejorar el
prompt de
autorefinamiento



Colocarlo en un
Gem de Gemini

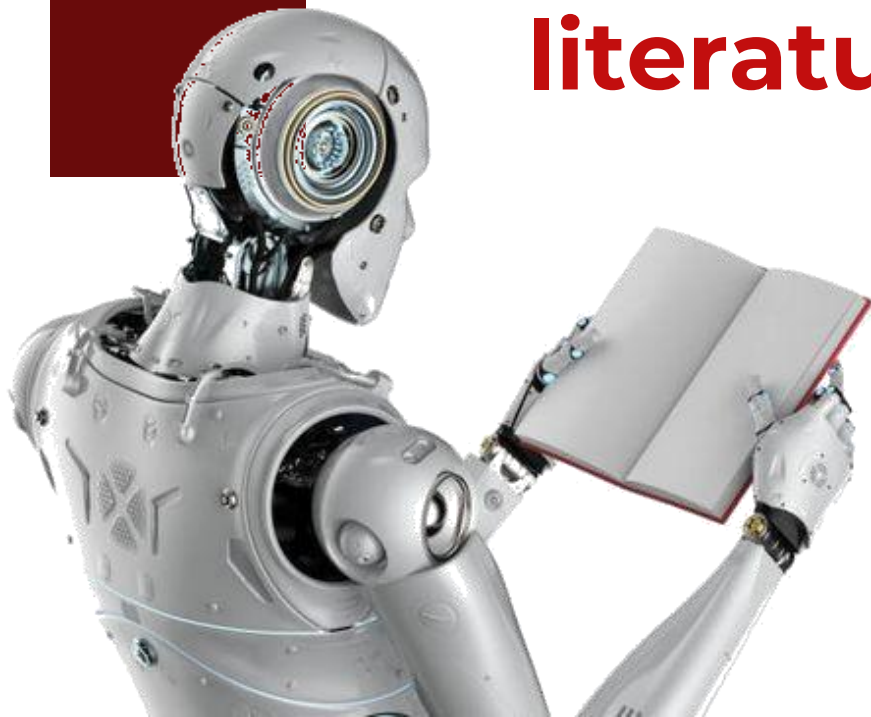


Prueba tu Gem
con un prompt
referente a tu
investigación

Taller I

02

Herramientas de IA para análisis de literatura científica



Materialles



Workshop_pii

02_LiteraturaCientifica

02 Tabla de contenidos



01

Diseño
metodológico



02

Revisión
Literaria



03

Investigación



04

Síntesis y
extracción



05

Redacción

Taller II

03

**Metodologías híbridas
para análisis de datos
asistido por IA**



Materials



Workshop_pii

03_Analisis