

# IA

Posibilidades, riesgos y buenas prácticas

# HOLA, SOY IMANOL

CONSULTOR DE ESTRATEGIA DIGITAL  
FORMADOR EN NUEVAS TECNOLOGÍAS

1. Introducción
2. Posibilidades
3. Riesgos
4. Buenas prácticas
5. Conclusiones

# índice



En esta presentación encontrarás contenido, tanto en formato texto como imagen, creado con herramientas de IA como: ChatGPT, Anthropic Claude, DALL-E 3 y RunwayML

# 01

## Introducción

Pongámonos en situación

- ChatGPT fue la gran **bomba de finales de 2022**.
- Ha **revolucionado** un montón de sectores.
- Inteligencia artificial que es capaz de **mantener conversaciones**.
- **Entiende** lo que le pedimos y nos da una respuesta coherente.
- Ha levantado mucho revuelo.
- **Primer puesto en Product Hunt 2022**

- Su irrupción no está exenta de **polémica**:
  - El robot conversacional ChatGPT aprueba el examen de la facultad de Derecho de Estados Unidos
  - Una famosa universidad francesa prohíbe el uso de ChatGPT o de la IA
  - Nueva York prohíbe ChatGPT en los colegios
  - Universidades australianas revisan sus métodos de evaluación
  - Los profesores de Málaga apuestan por la inteligencia artificial ChatGPT
  - Nuevos retos para el profesorado en la era de la inteligencia artificial

# 01

# Introducción



- Su irrupción no está exenta de **polémica**:
  - The Kindle Store has a prolific new author: ChatGPT
  - “Obligo a usar ChatGPT en mis clases”. Así es la irrupción inexorable de la nueva IA a las aulas
  - El sueño de China de tener su propio ChatGPT tiene un obstáculo: las severas restricciones comerciales de EEUU
  - Los cientos de miles de trabajadores en países pobres que hacen posible la existencia de sistemas de inteligencia artificial como ChatGPT (y por qué generan controversia)

# 01

# Introducción

- No se para de hablar de ChatGPT
  - ChatGPT llega a Slack: lo que puedes hacer con esta IA en tu día a día en el trabajo
  - ChatGPT ahora está disponible en Azure
  - DuckDuckGo está construyendo respuestas con IA en su buscador
  - Discord y OpenAI han llegado a un acuerdo para usar la IA de ChatGPT como bot especial en sus grupos.
  - Uno de los fundadores de OpenAI sale la compañía para evitar conflictos de intereses

01

Introducción

- No se para de hablar de ChatGPT
  - La Casa Blanca apoyará un hackeo masivo para poner a raya los chatbots de inteligencia artificial
  - Biden cita las tecnológicas en la Casa Blanca para intentar que no se descontrole la IA
  - La industria china de inteligencia artificial resiste pese al veto de EEUU a la exportación de chips
  - Filtraciones desde Google admiten estar perdiendo la carrera en inteligencia artificial: "Un tercero se nos come el almuerzo"
  - Amazon quiere que Alexa tenga Inteligencia Artificial a la ChatGPT: el nuevo Bing parece haber dejado en coma a los asistentes de voz

01

Introducción

- No se para de hablar de ChatGPT
  - Petición de parar 6 meses el desarrollo de IA.
  - AEPD abre investigación sobre ChatGPT.
  - ChatGPT baneado en Italia
  - Google deja a la UE fuera de su nueva inteligencia artificial
  - Luz verde para la nueva ley de inteligencia artificial de Europa: así te va a afectar

# 01

# Introducción

- No se para de hablar de ChatGPT
  - La Unesco pide reglas sobre el uso de la IA en las escuelas
  - Microsoft anuncia un nuevo compromiso de derechos de autor de Copilot para los clientes
  - Microsoft colabora con Paige para construir el modelo de imágenes para luchar contra el cáncer más grande.
  - A casi un año de ChatGPT, ¿qué piensan los directivos escolares de la inteligencia artificial?
  - La Agencia Española de Inteligencia Artificial estará funcionando en 3 meses como máximo

# 01

# Introducción

## CRONOLOGÍA

- **Junio 2018:** OpenAI anuncia GPT-1
- **Febrero 2019:** OpenAI anuncia GPT-2
- **Mayo 2020:** Primer paper sobre GPT-3
- **Junio 2020:** Beta privada API GPT-3
- **Septiembre 2020:** GPT-3 licenciada a Microsoft
- **Noviembre 2021:** API GPT-3 abierta al público

01

Introducción

# CRONOLOGÍA

- **Enero 2022:** Salida de InstructGPT
- **Septiembre 2022:** 66% bajada de precio davinci model
- **Septiembre 2022:** Se anuncia Whisper
- **Noviembre 2022:** GPT-3.5 se amplia con text-davinci-003
- **Noviembre 2022:** Se anuncia ChatGPT

01

Introducción

## CRONOLOGÍA

- **Febrero 2023:** Se presenta Bing Chat
- **Marzo 2023:** Se abre la API de ChatGPT
- **Marzo 2023:** Se anuncia GPT-4
- **Marzo 2023:** Se anuncia Microsoft 365 Copilot
- **Marzo 2023:** Plugins en ChatGPT
- **Mayo 2023:** ChatGPT app para iOS



## CRONOLOGÍA

- **Julio 2023:** Custom Instructions en ChatGPT
- **Julio 2023:** App para Android
- **Agosto 2023:** ChatGPT Enterprise
- **Septiembre 2023:** Anuncio de DALL·E 3
- **Septiembre 2023:** Anuncio de la multimodalidad

01

Introducción

# ChatGPT

En cifras

# 300

AÑOS DE INFORMACIÓN EN SU MODELO ORIGINAL

[https://docs.google.com/presentation/d/17b\\_ocq-GL5lhV\\_bYSShzUgxL02mtWDoiw9xEroJ5m3Q/edit#slide=id.g1b8e0b333f6\\_0\\_13](https://docs.google.com/presentation/d/17b_ocq-GL5lhV_bYSShzUgxL02mtWDoiw9xEroJ5m3Q/edit#slide=id.g1b8e0b333f6_0_13)

01

Introducción

**\$5.000.000**

COSTE DE DESARROLLO APROXIMADO

[https://docs.google.com/presentation/d/17b\\_ocq-GL5lhV\\_bYSShzUgxL02mtWDoiw9xEroJ5m3Q/edit#slide=id.g1b8e0b333f6\\_0\\_13](https://docs.google.com/presentation/d/17b_ocq-GL5lhV_bYSShzUgxL02mtWDoiw9xEroJ5m3Q/edit#slide=id.g1b8e0b333f6_0_13)

01

Introducción

**\$3.000.000**

AL MES PARA MANTENERLO DISPONIBLE

<https://www.digitaltrends.com/computing/how-to-use-openai-chatgpt-text-generation-chatbot/>

01

Introducción

# \$10.000M

INVERSIÓN DE MICROSOFT EN OPENAI

<https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai?sref=10INAhZ9>

01

Introducción

# 100.000.000

USUARIOS ÚNICOS MENSUALES ACTIVOS

<https://nerdynav.com/chatgpt-statistics/#chatgpt-user-growth>

01

Introducción



[https://twitter.com/whoiskatrin/status/1600421531212865536?s=12&t=LQIBtblGra0LnjawoHHI3g\]\(https://twitter.com/whoiskatrin/status/1600421531212865536?s=12&t=LQIBtblGra0LnjawoHHI3g\)](https://twitter.com/whoiskatrin/status/1600421531212865536?s=12&t=LQIBtblGra0LnjawoHHI3g](https://twitter.com/whoiskatrin/status/1600421531212865536?s=12&t=LQIBtblGra0LnjawoHHI3g))

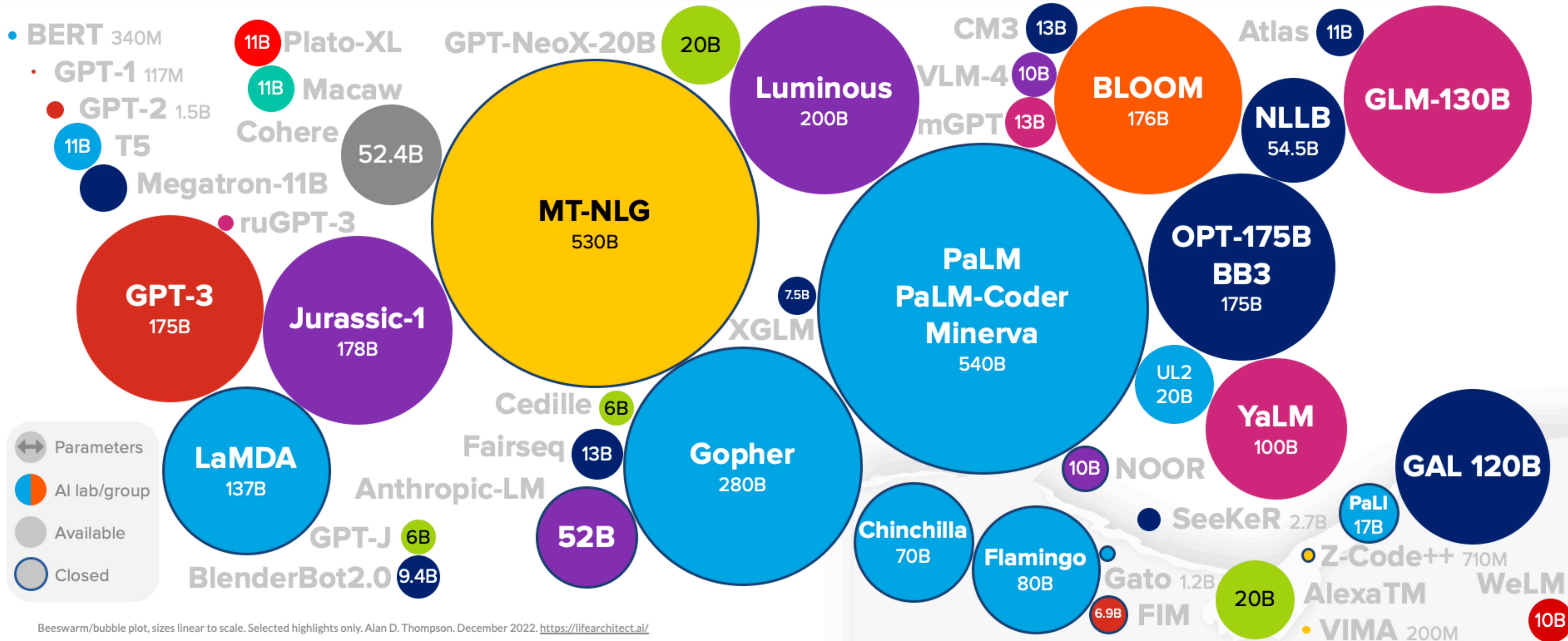
# 01 Introducción



## PRÓXIMOS PASOS

- Microsoft 365 Copilot
- Windows Copilot
- Google Duet
- GPT4

# LANGUAGE MODEL SIZES TO DEC/2022

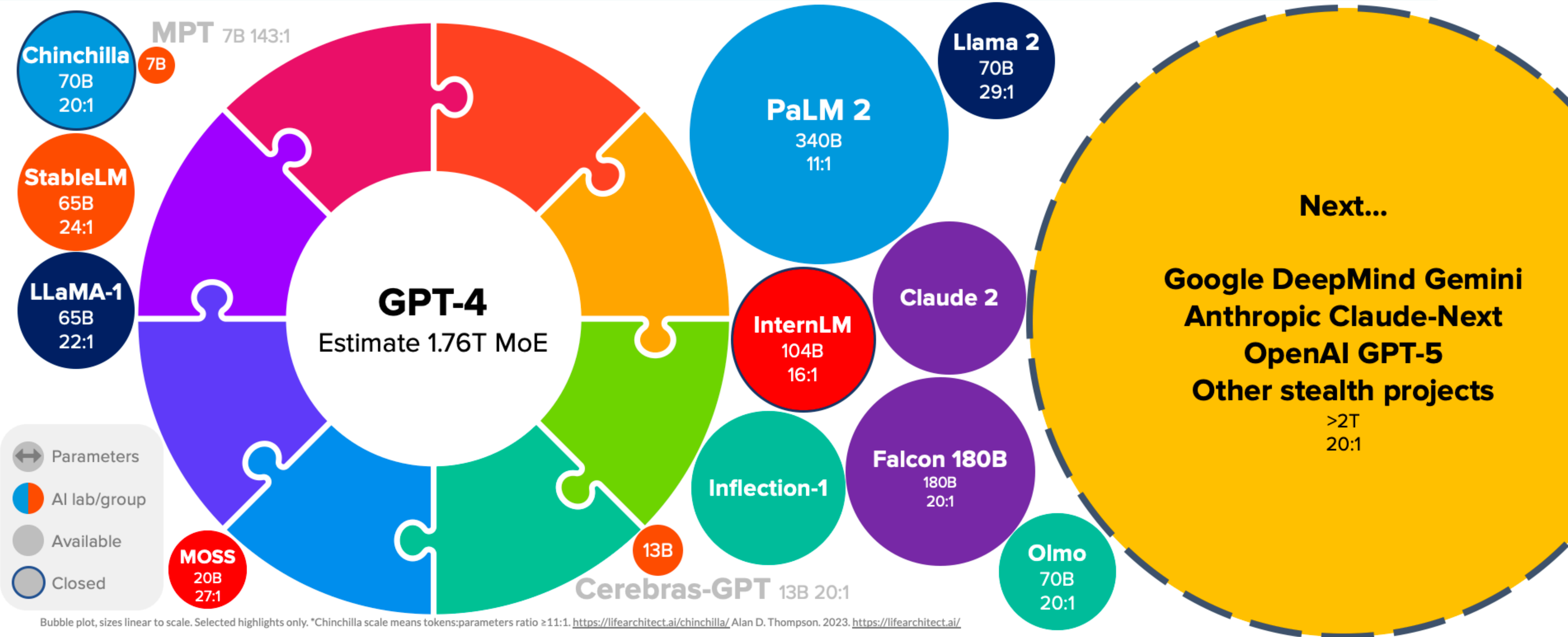


Beeswarm/bubble plot, sizes linear to scale. Selected highlights only. Alan D. Thompson. December 2022. <https://lifearchitect.ai/>



# 2023-2024 OPTIMAL LANGUAGE MODELS

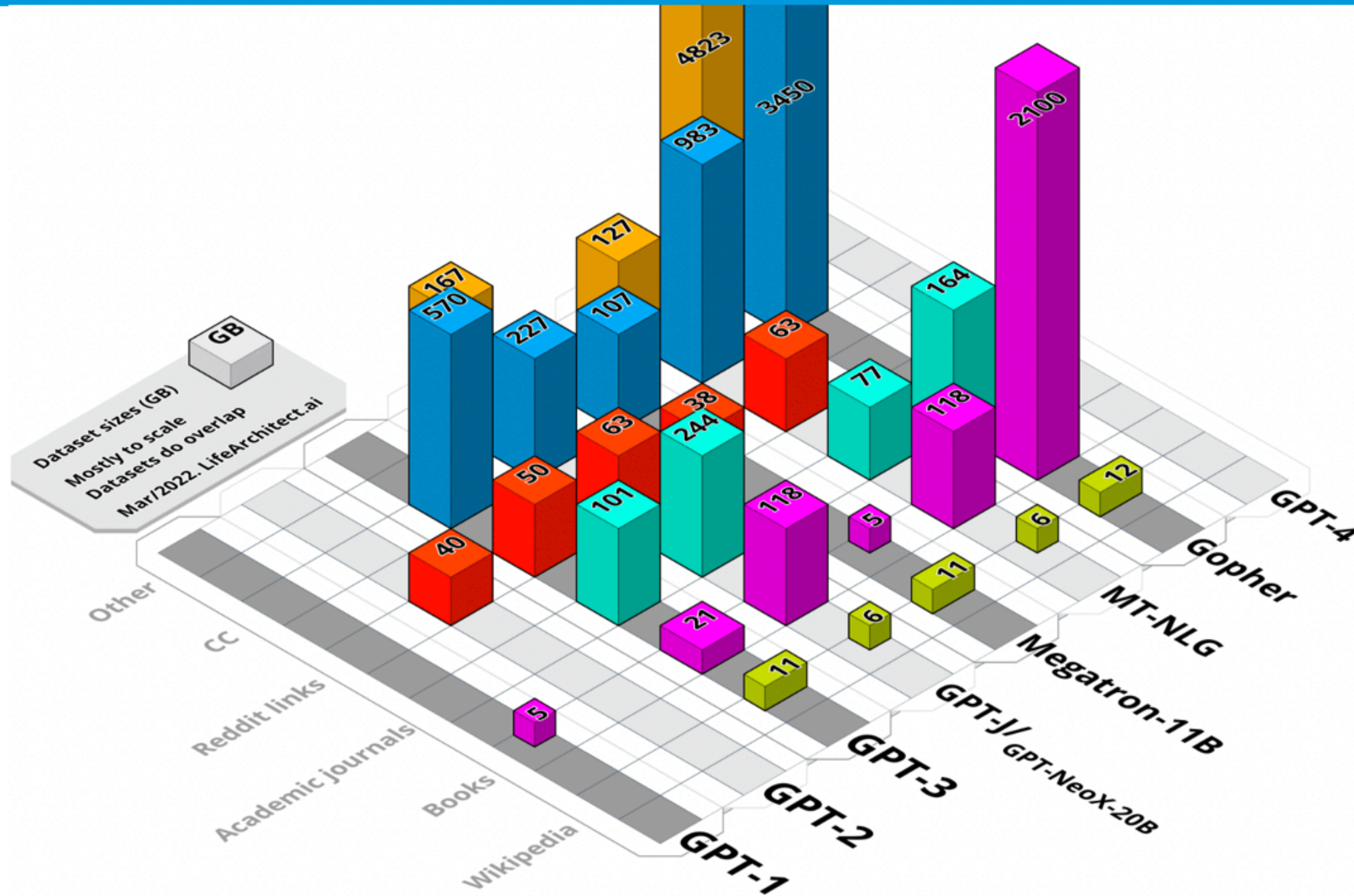
SEP/  
2023



Bubble plot, sizes linear to scale. Selected highlights only. \*Chinchilla scale means tokens:parameters ratio  $\geq 11:1$ . <https://lifearchitect.ai/chinchilla/> Alan D. Thompson. 2023. <https://lifearchitect.ai/>



# 2022 WHAT'S IN MY AI? – ALT VIEW



Google Patents.....	0.48%
The New York Times.....	0.06%
Los Angeles Times.....	0.06%
The Guardian.....	0.06%
Public Library of Science..	0.06%
Forbes.....	0.05%
Huffington Post.....	0.05%
Patents.com.....	0.05%
Scribd.....	0.04%
Other.....	99.09%

**Common Crawl**

Google.....	3.4%
Archive.....	1.3%
Blogspot.....	1.0%
GitHub.....	0.9%
The New York Times.....	0.7%
Wordpress.....	0.7%
Washington Post.....	0.7%
Wikia.....	0.7%
BBC.....	0.7%
Other.....	89.9%

**Reddit links**

Biography.....	27.8%
Geography.....	17.7%
Culture and Arts.....	15.8%
History.....	9.9%
Biology, Health, Medicine.....	7.8%
Sports.....	6.5%
Business.....	4.8%
Other society.....	4.4%
Science & Math.....	3.5%
Education.....	1.8%

**English Wikipedia**

Romance.....	26.1%
Fantasy.....	13.6%
Science Fiction.....	7.5%
New Adult.....	6.9%
Young Adult.....	6.8%
Thriller.....	5.9%
Mystery.....	5.6%
Vampires.....	5.4%
Horror.....	4.1%
Other.....	18.0%

**BookCorpus (GPT-1 only)**



# 84%

More truthful



## HHH: Helpful, honest, harmless

New alignment objective to be useful, truthful, and careful

# InstructGPT

 OpenAI

# 58%

Less hallucinative

# 1.5 years

More knowledge

To June 2021



# 1.9x

Larger context window  
(from 2,048 to 4,000 tokens)



OpenAI InstructGPT one-pager. Compared to GPT-3. Sources: <https://openai.com/blog/instruction-following/> and <https://beta.openai.com/docs/engines>. Alan D. Thompson. April 2022. <https://lifearchitct.ai/>



02

# Posibilidades

Capacidad de conversación y aprendizaje

- ChatGPT puede utilizarse para un **sinfín** de **propósitos**.
- No deja de ser un **asistente**, al que le pedimos lo que necesitamos y nos da un resultado.
- De lo **precisos** que seamos en la petición, dependerá la calidad del resultado obtenido.
- Por lo tanto, puede utilizarse para casi **cualquier cosa** que podamos necesitar en nuestro día a día.

02

Posibilidades

*Se llama **prompt** al carácter o conjunto de caracteres que se muestran en una línea de comandos para indicar que está a la espera de órdenes. Este puede variar dependiendo del intérprete de comandos y suele ser configurable.*

02

Posibilidades



- Hay que tener siempre en cuenta que **no nos podemos fiar al 100%** del contenido generado.
- Es **nuestra responsabilidad** cerciorarnos que el resultado es de calidad.
- Tema especialmente sensibles en según qué casos:
  - Salud.
  - Aspectos legales.

02

Posibilidades

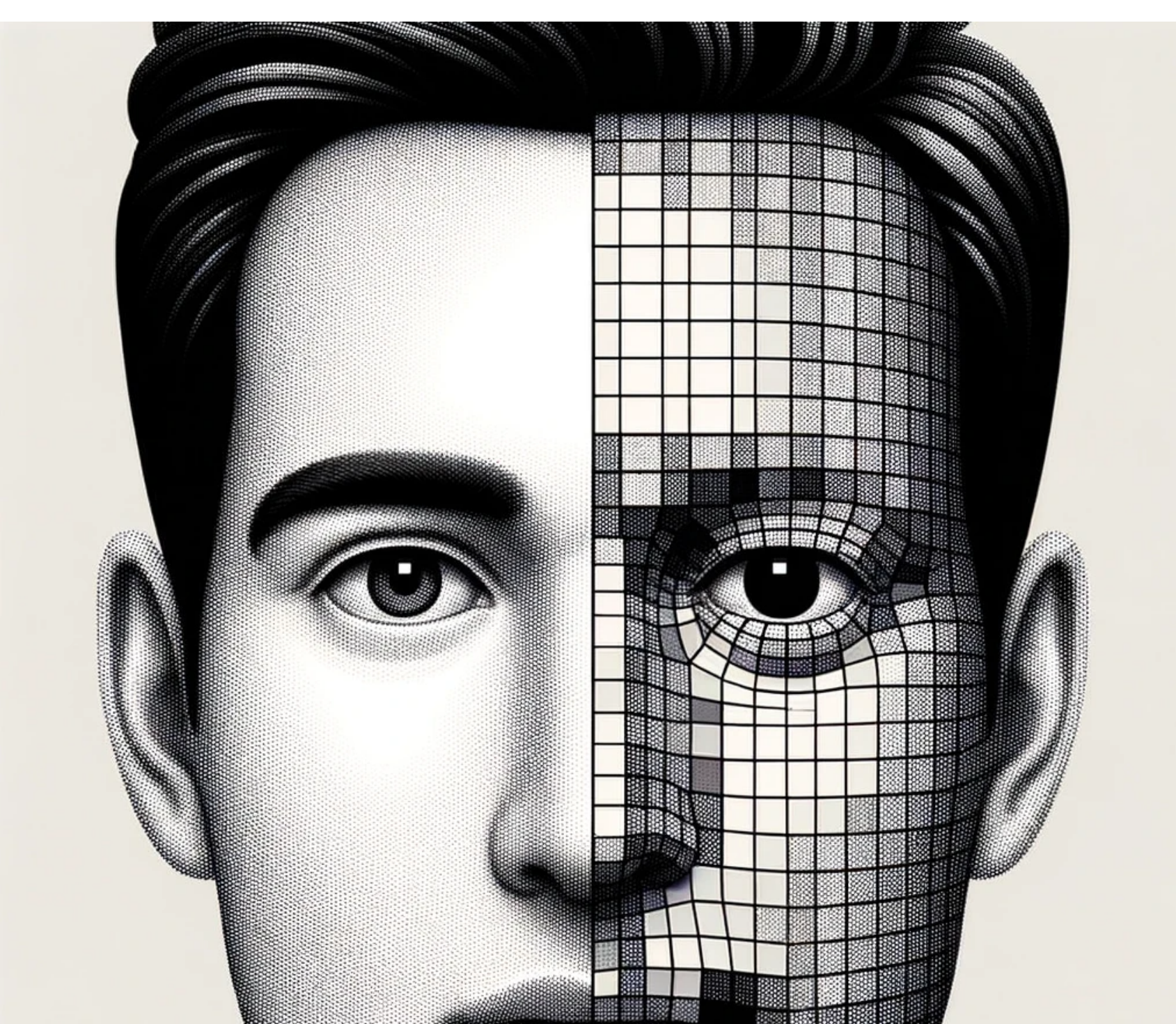
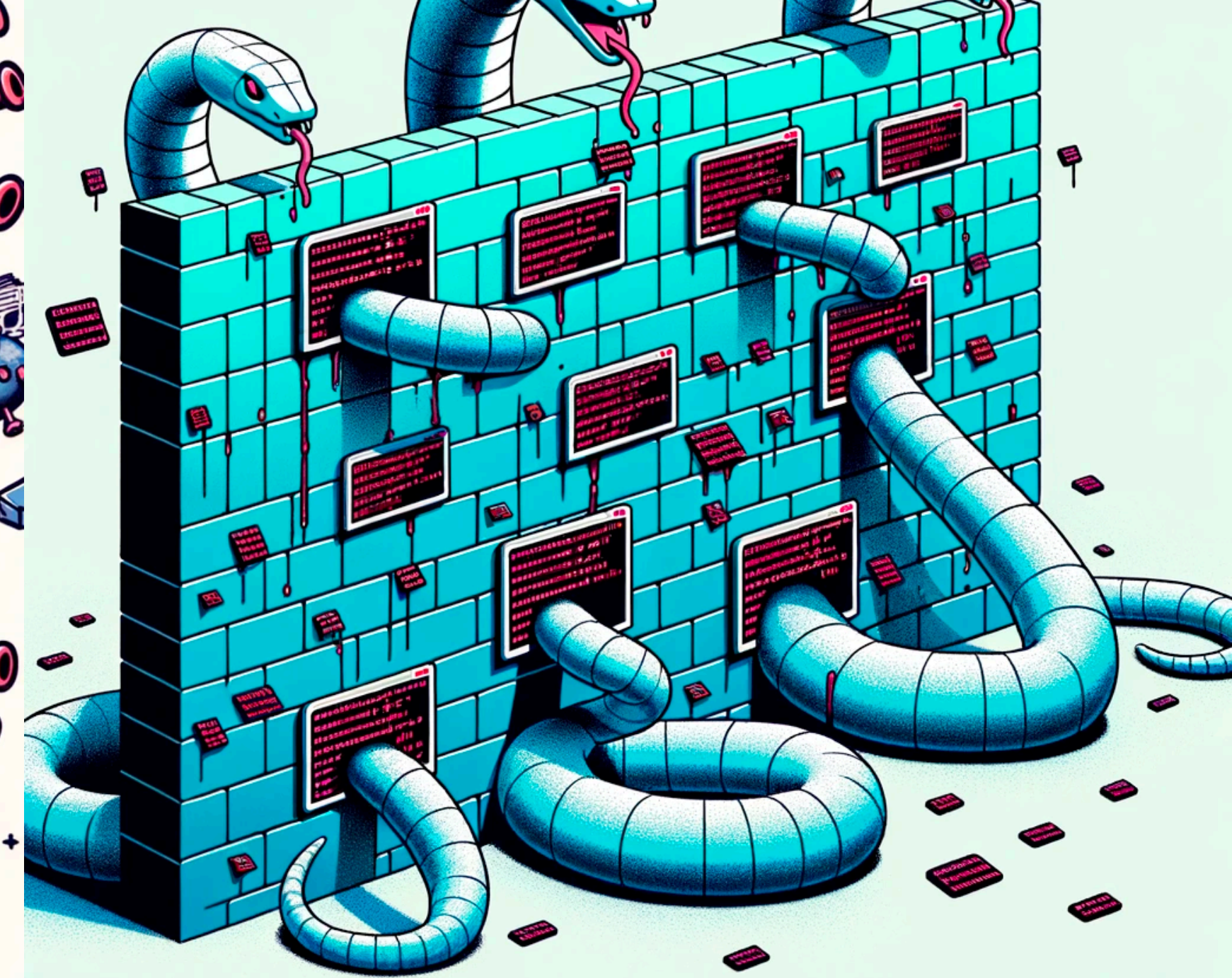
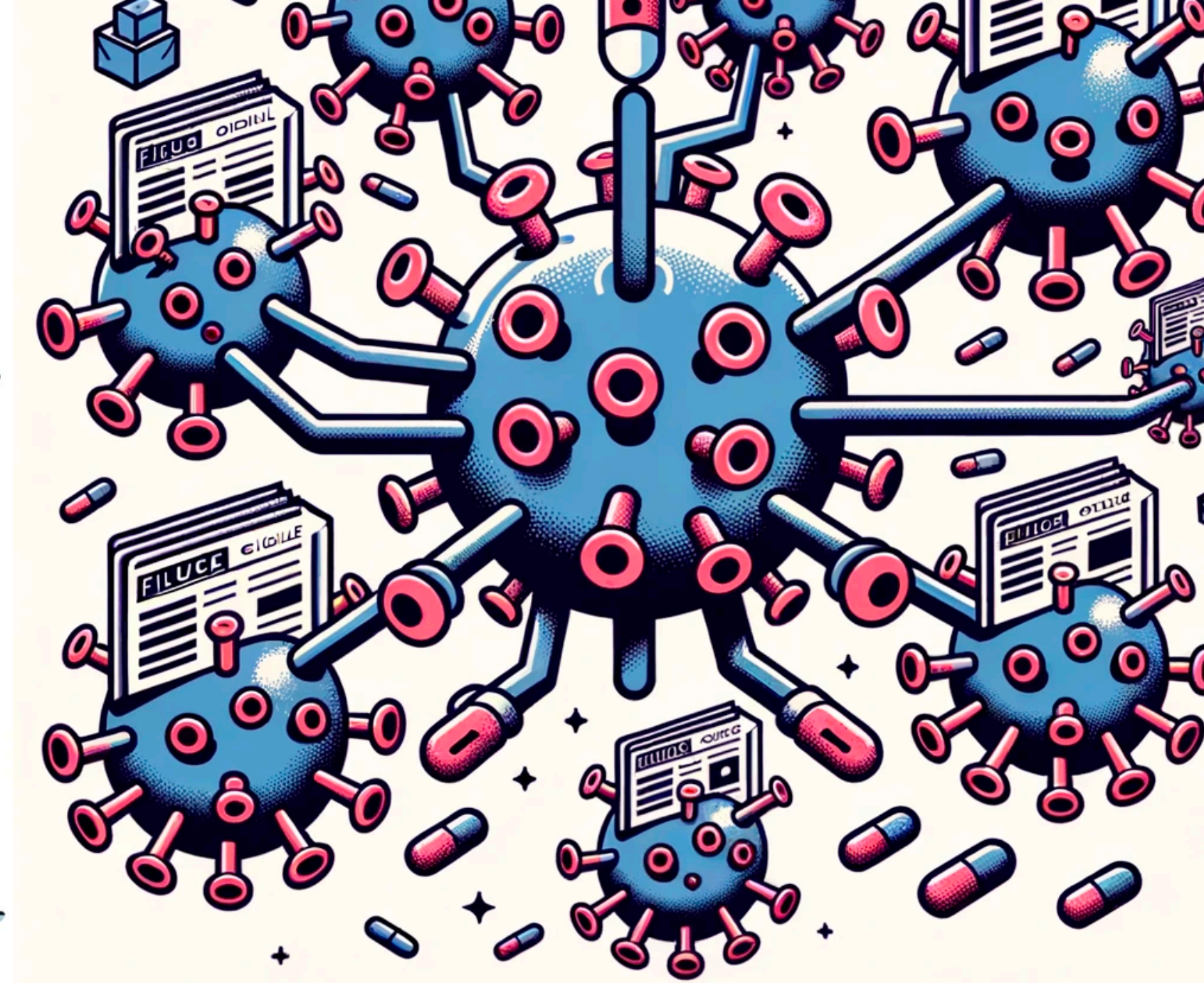
- Redacción en cualquier formato.
- Resúmenes.
- Traducciones.
- Ampliar textos.
- Explicar de otra manera.
- Análisis de textos.
- Código de programación.

Y mucho más...

03

# Riesgos

Y amenazas



- **Privacidad** de los datos y **confidencialidad** de la información.
- **Manipulación** de información y **desinformación**.
- **Vulnerabilidad** frente a **ataques** cibernéticos.
- **Suplantación** de identidad.
- **Phishing**.
- Perpetuación de **sesgos**.

03

Riesgos







A close-up, slightly blurred photograph of a man's face, looking downwards and to the left. He has short, dark hair and a light beard. The background is a solid pink color. The image is framed by blue L-shaped corner brackets in the top-left and bottom-right corners.

**¡HABLANDO  
EN INGLÉS!**

The logo for T Sports, featuring a stylized 'T' in a blue square followed by the word 'SPORTS' in a blue, outlined font.

**T SPORTS**



04

# Buenas prácticas

Y medidas para limitar riesgos

1. Sé **transparente** sobre el uso de IA generativa.

Si generas contenido con IA, comunícalo. Esto crea confianza y evita confusión.

2. Ten **cuidado** con **sesgos** y desinformación.

Evalúa críticamente el contenido generado por IA para detectar posibles sesgos o información falsa. Entrena los modelos con datos diversos y de buena calidad.

3. **Respetar** los **derechos** de autor y la **propiedad** intelectual.

No reclames la autoría de contenido generado por IA, y ten cuidado de no infringir derechos de autor.

4. **Supervisa** el **contenido** generado.

Revisa manualmente el contenido de IA antes de publicarlo para detectar errores, contenido inapropiado o posible plagio.

5. Sé **responsable** con **datos** sensibles.

Ten cuidado al generar contenido relacionado con finanzas, salud o identidades personales. Sigue las leyes de privacidad de datos.

6. Proporciona **contexto** y **atribución**.

Si publicas contenido de IA, indica claramente que fue generado por IA y proporciona el contexto adecuado para evitar malentendidos.

7. Ten un **proceso** de **gobernanza**.

Establece pautas claras para el uso ético y responsable de herramientas de IA generativa dentro de su organización.

8. **Actualiza modelos** de IA regularmente.

Los sesgos y errores pueden reducirse entrenando los modelos con nuevos datos de forma continua.

04

Buenas prácticas

# 05

## Conclusiones

Y consideraciones finales

- El **mundo**, tal y como lo conocíamos, **ha cambiado**.
- Es un **cambio profundo y transversal**.
- Hay que **adaptarse**, aprender cómo funcionan estas tecnologías.
- Hacer un **uso responsable**.
- **Verificar** información.



# IA GENERATIVA

Máster Online

[iagenerativa.com](http://iagenerativa.com)

# imanol terán.

[hola@imanolteran.com](mailto:hola@imanolteran.com)

**Eskerrik asko**