

Como Gestionar correctamente las IAs en las AAPP



Índice

- Sobre nosotros	3
- Café ¿con leche?	4
- La IA, ¿pero qué IA?	6
- IA generativa	7
- De gestionar a gestionar correctamente	10
- ¿El final del principio?	15

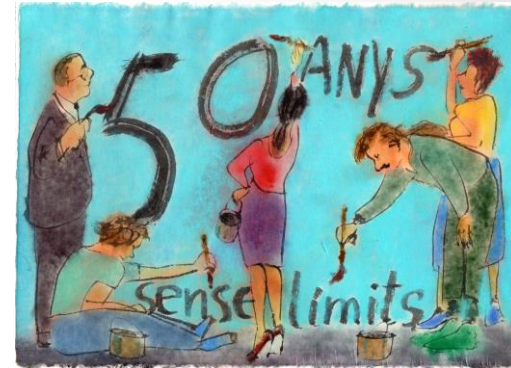
TODOS los datos citados en la presentación tienen sus fuentes en el de apartado referencias

Sobre nosotros...

- **Contexto:**

➤ UPC, Universitat Politècnica de Catalunya, <https://www.upc.edu>

- Institución pública
- Investigación y educación superior
- Ingeniería, arquitectura, ciencias y tecnología
- Desde 1971



➤ /rdlab, Research and Development Lab, <https://rdlab.cs.upc.edu>

- Dedicado al soporte TIC de grupos de investigación
- Colaboramos con +20 grupos y 600 investigadores (UPC + colaboradores)
- Proyectos nacionales/europeos y de transferencia de tecnología (+65)
- Especializados en HPC y servicios Cloud para la investigación
- Gestionamos más de 200 servidores físicos y 300TBytes de almacenamiento
- Desde 2010, Computer Science Department (<https://www.cs.upc.edu>)

Café con ¿leche?



Pero ¿qué leche!

MILK NUTRITION GUIDE

NUTRITION COMPARISON FOR 8oz CUP | MORE AT CHEATDAYDESIGN.COM

 <p>Skim 80 Calories 0g Fat 12g Carbs 11g Sugar 8g Protein</p>	 <p>1% 100 Calories 2.5g Fat 12g Carbs 12g Sugar 8g Protein</p>	 <p>2% 120 Calories 5g Fat 12g Carbs 12g Sugar 8g Protein</p>	 <p>Whole 150 Calories 8g Fat 12g Carbs 12g Sugar 8g Protein</p>	 <p>Soy 110 Calories 4.5g Fat 9g Carbs 6g Sugar 8g Protein</p>
 <p>Coconut 70 Calories 4.5g Fat 6g Carbs 5g Sugar 0g Protein</p>	 <p>Coconut (Unsweetened) 40 Calories 4g Fat 1g Carbs 0g Sugar 0g Protein</p>	 <p>Almond/Cashew (Unsweetened) 25 Calories 2g Fat 1g Carbs 0g Sugar 1g Protein</p>	 <p>Almond (Sweetened) 60 Calories 2.5g Fat 8g Carbs 7g Sugar 1g Protein</p>	 <p>Silk Protein 130 Calories 8g Fat 3g Carbs 2g Sugar 10g Protein</p>
 <p>Oat 100 Calories 6g Fat 11g Carbs 8g Sugar 2g Protein</p>	 <p>Rice 120 Calories 2.5g Fat 23g Carbs 10g Sugar 1g Protein</p>	 <p>Flax 50 Calories 2.5g Fat 7g Carbs 7g Sugar 0g Protein</p>	 <p>Walnut 120 Calories 11g Fat 1g Carbs 0g Sugar 3g Protein</p>	 <p>Hemp 130 Calories 6g Fat 14g Carbs 4g Sugar 3g Protein</p>
 <p>Macadamia 70 Calories 4.5g Fat 7g Carbs 6g Sugar 1g Protein</p>	 <p>Hazelnut 110 Calories 3.5g Fat 19g Carbs 14g Sugar 2g Protein</p>	 <p>Pea 90 Calories 4.5g Fat 6g Carbs 5g Sugar 8g Protein</p>	 <p>Pistachio 100 Calories 8g Fat 4g Carbs 1g Sugar 4g Protein</p>	 <p>Goat 140 Calories 7g Fat 11g Carbs 11g Sugar 8g Protein</p>

THIS IS AN ALMOND. →



IT IS NOT A MAMMAL.

ONLY MAMMALS PRODUCE MILK.

SO I REALLY DON'T KNOW WHAT YOU'VE BEEN PUTTING IN YOUR COFFEE...

...BUT IT AIN'T MILK.

cheatdaydesign.com

IA...¿pero qué IA?



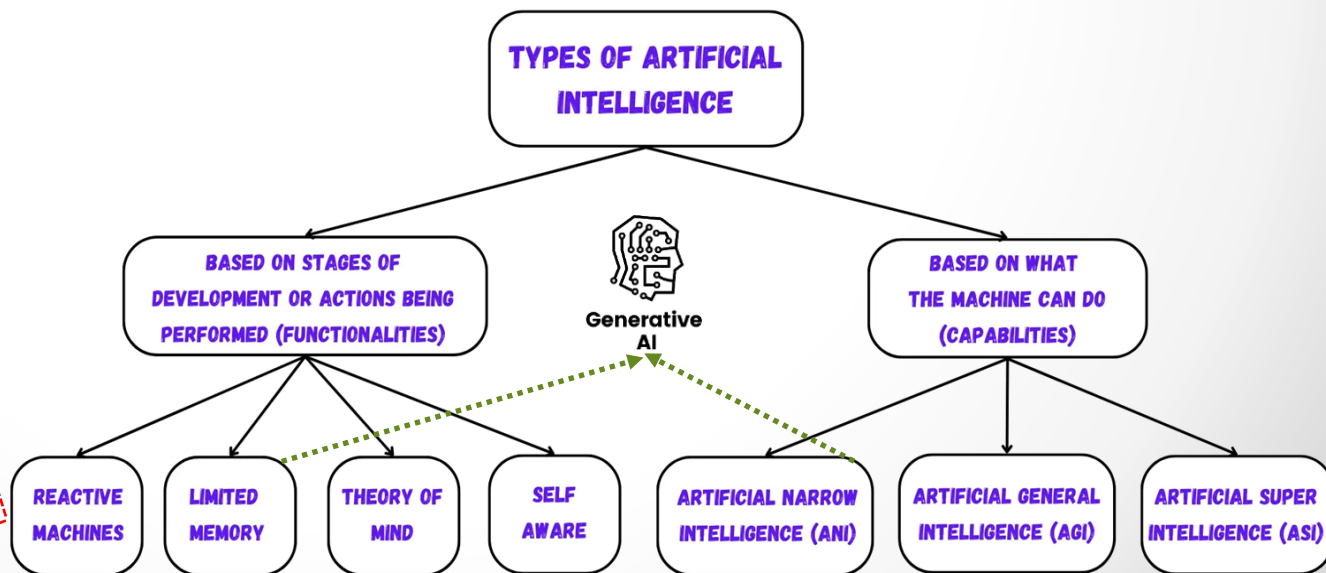
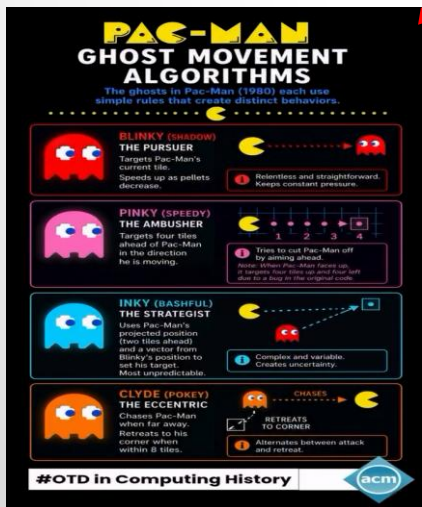
¿Inteligencia? ¿Artificial?



1. adj. Hecho o producido por el ser humano.
2. adj. No natural, falso.

¡Hablemos con propiedad! (el lenguaje crea realidades):

- Me lo dijo la **Inteligencia**
- Me lo dijo la **Inteligencia Artificial**
- Me lo dijo la **Inteligencia Artificial generativa / modelo LLM**



IA generativa II: Generative Pre-trained Transformer

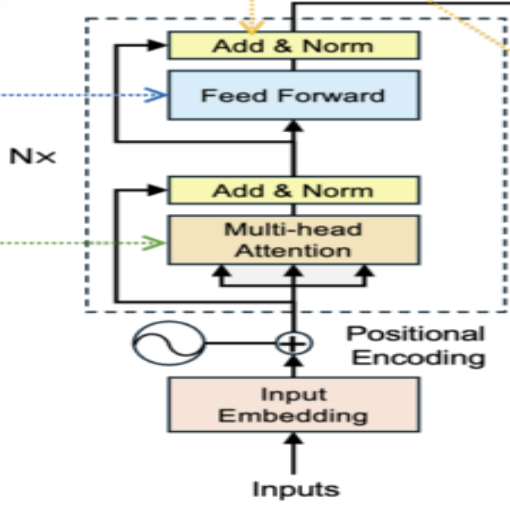
2017 Google – Attention is all you need

2018 OpenAI – GPT-1

Feed Forward Network
Perform matrix multiplication between the input and the weight matrix, followed by adding the bias vector.

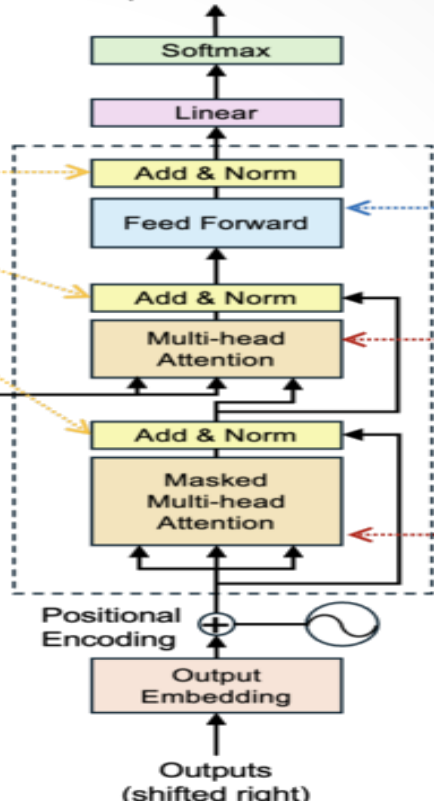
Encoder Self-attention
Produce an encoded representation for each word in the input sequence.
Queries, keys, and values input from encoder states.

Residual connections and layer normalization



BERT-style
(Encoder-only)

Output Probabilities



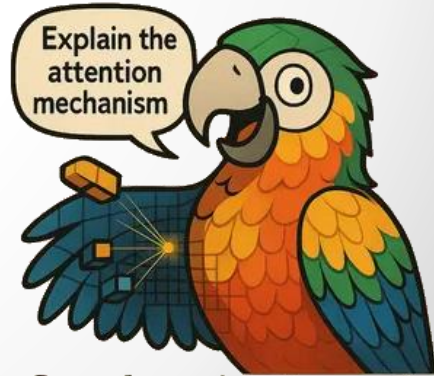
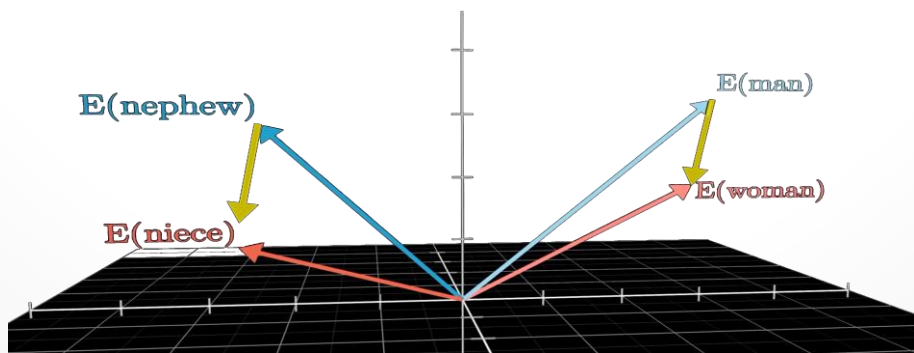
Feed Forward Network
Perform matrix multiplication between the input and the weight matrix, followed by adding the bias vector.

Decoder-Encoder Self-attention
Produce an encoded representation of both the target sequence and the input sequence.
queries from decoder states; keys and values from encoder states

Decoder Self-attention
Produce an encoded representation for each word in the target sequence.
queries, keys, and values input from decoder states.

GPT-style
(Decoder-only)

$$E(\text{niece}) - E(\text{nephew}) \approx E(\text{woman}) - E(\text{man})$$



Stochastic Parrot

IA generativa III: ¿alucinaciones? ¿no sé?



Camillo Miola (Biacca) – L'Oracolo

De gestionar a gestionar correctamente



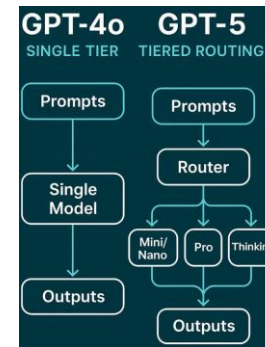
Opinión: Juicio o valoración que se forma una persona respecto de algo o de alguien.

Criterio: Norma para conocer la verdad.

- **Gestionar las IAs en las AAPP (holística)**

- Sabemos cómo funcionan
- Sabemos que son sistemas estocásticos (pero **NO** loros)
- Sabemos que (“*actualmente*”) son sistemas de IA **generativa** (≠general)

E
S
T
R
A
T
E
G
I
A



- **Gestionar correctamente las IAs en las AAPP (analítica)**

- Perspectiva tecnológica
- Perspectiva legal
- Perspectiva ética

“...because without (proper) data, you are just another person with an opinion.”

Featuring William Edwards Deming

De gestionar a gestionar correctamente II

• Perspectiva tecnológica

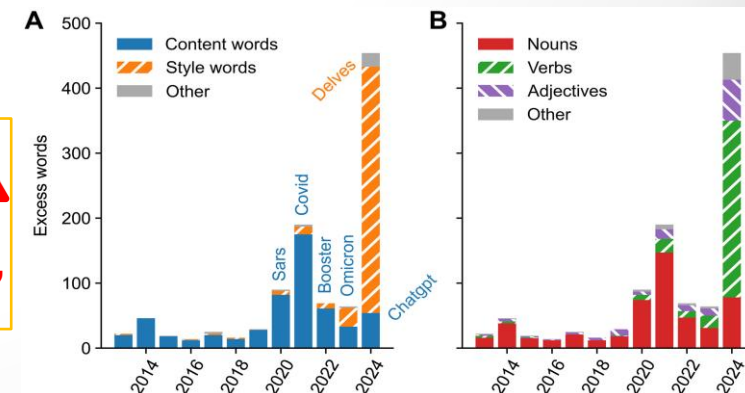
- ¿Con que conjunto de datos ha sido **entrenado**? ¿Y **personalizado**?
 - Idioma: % respecto inglés, Valenciano-Mallorquín-Catalán
 - Contenido: The Anarchist Cookbook / Wikipedia / IA slop
- Si personalizamos/enriquecemos el modelo LLM:
 - Crear **ad hoc** un conjunto de entrenamiento / control / **pruebas** (>> 50%)
 - ¿Debe estar abierto? (toolboxes) ↔ Sistemas salvaguardia (Guardrails)
- Limitaciones en los ámbitos de aplicación:
 - Extremadamente buenos en traducciones, resúmenes, Noob (inexpertos)
 - Discutibles en predicciones, hechos y razonamiento



§ Efecto “Eliza” (1966) / Dr. Sbaitsso

§ MrGoxx / CNMV: LLM and stock investing

§ Science: “**Delving into LLM-assisted writing in biomedical...**”



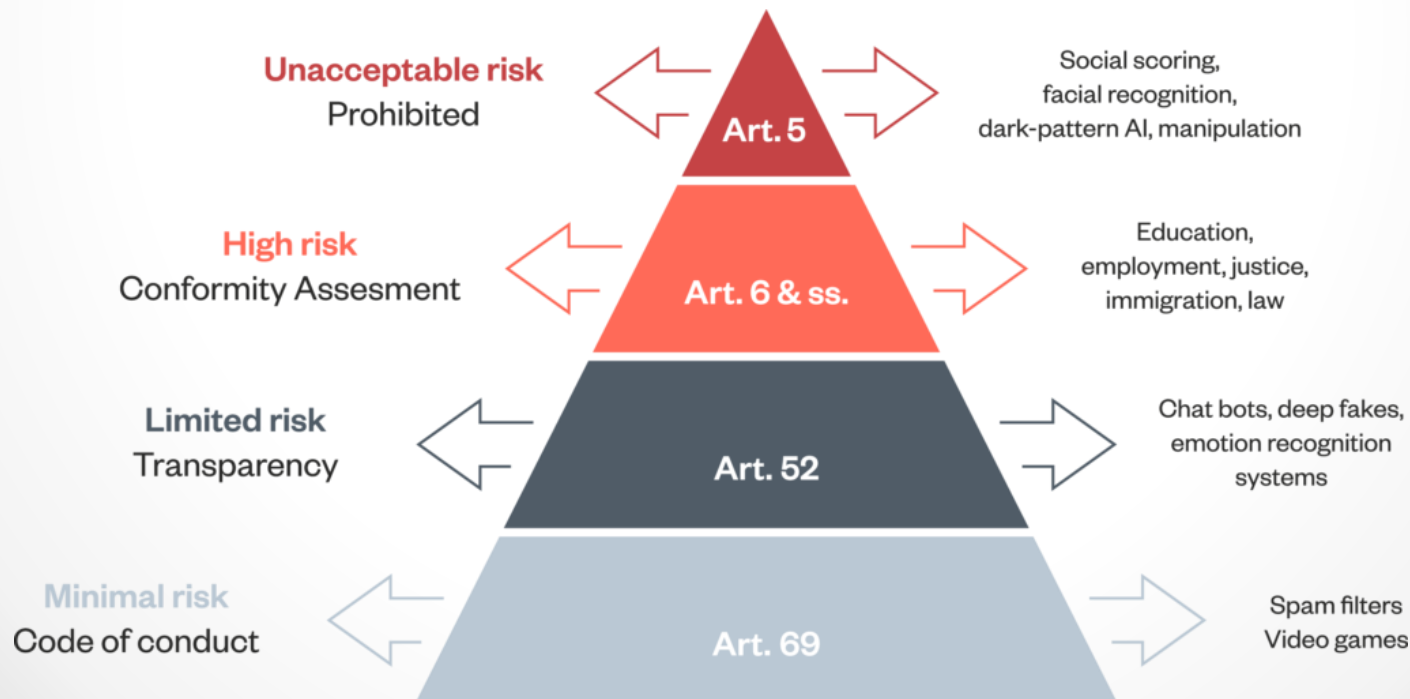
De gestionar a gestionar correctamente III

- **Perspectiva legal (datos, conversaciones, resultados y usos)**

➤ GDPR / LOPD (DMP, consentimiento informado, análisis de riesgos, evaluación de impacto)

Article 3(1) ensures that the GDPR applies to the processing by a controller or processor carried out in the context of the activities of an establishment of that controller or processor in the Union, regardless of the actual place of the processing.

➤ ARTIFICIAL INTELLIGENCE ACT / RIA



De gestionar a gestionar correctamente IV

- **Perspectiva ética (AAPP) – CE /CEI /CEIm**

- Valores de la institución ↔ Doble uso / Previsión vs Predicción
- Sesgos (falta real de datos) vs ideología (sobrerrepresentación)
- Costes y alternativas: Medioambiental, de oportunidad, de mantenimiento (Coherencia: Calculadora-LLM vs Hojas de cálculo / Uso de “por favor”)
- Digitalización forzosa (Exclusión digital, nuevos **token**-problemas)

Token Maxxing or **Token Maxing** is a metric used in an attempt to track productivity in the workplace especially for those using [generative] AI.



Previsión vs Predicción

¿Qué alumnos de primer curso necesitaran ayuda?

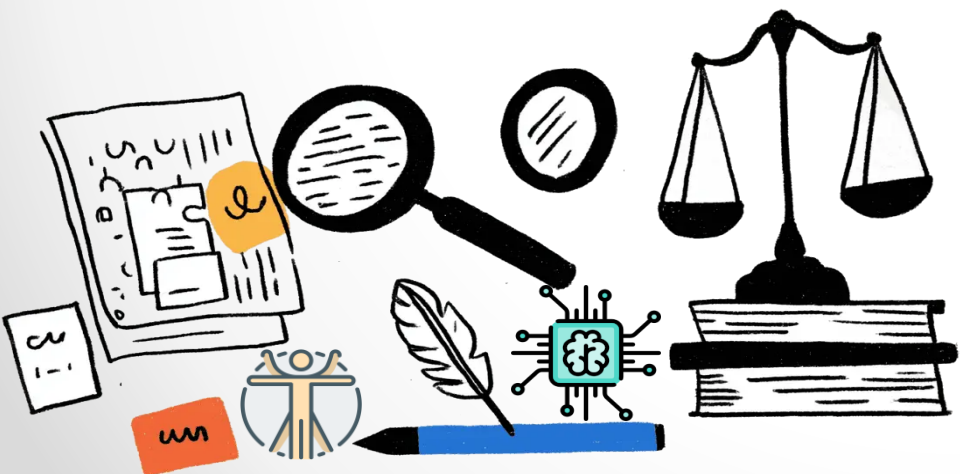


De gestionar a gestionar correctamente V

- ¿Cómo implementarlo?

Para dar el salto de "gestionar" (hacer que las cosas funcionen) a "gestionar correctamente" (hacerlas con excelencia y eficiencia), la clave está en pasar de la reacción a la estrategia.

- **Ámbito:** Proyectos transversales en AAPP + Investigación/TT => CE, CEI, CEIm...
- **No** hay un diseño único → Personalizarlo respetando la separación de los 3 ámbitos
- **Nuevos Roles:** Chief AI Officer (CAIO) / DPD / Ethics Mentor
- **Transparencia:** Quorum unánime (cada uno vota binariamente sólo su ámbito)



Proyecto: XXX

	Sí	No
Perspectiva Tecnológica	✓	
Perspectiva Legal		✗
Perspectiva Ética	✓	

¿El final del principio?

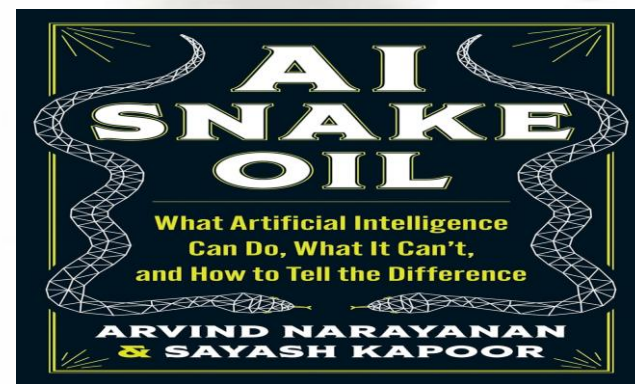
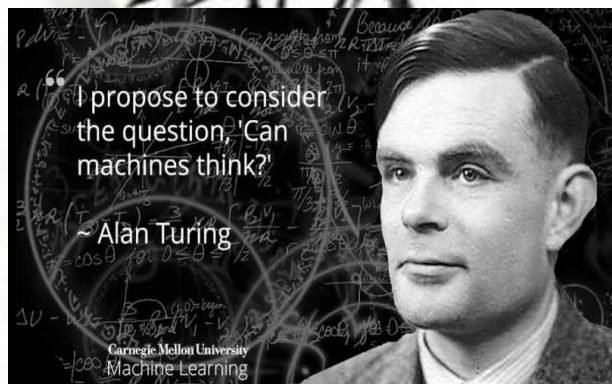


- Una cita, una reflexión y un libro:



Parte de la inhumanidad de las computadoras es que, una vez que están programadas competentemente y trabajando correctamente, son completamente honestas.

--- Isaac Asimov %



Referencias

- [www1] [https://commons.wikimedia.org/wiki/File:Camillo_Miola_\(Biacca\)_-The_Oracle_-_72.PA.32_-_J._Paul_Getty_Museum.jpg](https://commons.wikimedia.org/wiki/File:Camillo_Miola_(Biacca)_-The_Oracle_-_72.PA.32_-_J._Paul_Getty_Museum.jpg)
- [www2] <https://www.youtube.com/watch?v=7xTGNNLPyMI>
- [www3] <https://drive.google.com/file/d/1EZh5hNDzxMMY05uLhVryk061QYQGTXiN/view>
- [www4] <https://www.youtube.com/watch?v=Pnd8bCJ4Z3A>
- [www5] <https://www.youtube.com/watch?v=eMlx5fFNoYc>
- [www6] <https://www.youtube.com/watch?v=U2hZFMVNSE0>
- [www7] https://i.blogs.es/421374/cafe-con-leche2/1366_2000.jpg
- [www8] <https://cheatdaydesign.com/milk-nutrition-guide/>
- [www9] <https://tea-ink.com/products/almond-milk>
- [www10] <https://shriramsivanandhan.medium.com/artificial-intelligence-machine-learning-deep-learning-an-overview-86343909f990>
- [www11] <https://arcitech.ai/types-of-artificial-intelligence/>
- [www12] <https://intersog.co.il/blog/technologies/artificial-intelligence-in-a-nutshell-types-principles-and-history/>
- [www13] <https://www.forbes.com/sites/cognitiveworld/2019/06/19/7-types-of-artificial-intelligence/>
- [www14] <https://www.bmc.com/blogs/artificial-intelligence-types/>
- [www15] https://www.researchgate.net/figure/The-encoder-decoder-structure-in-LLMs-adjusted-from-Vaswani-et-al-74_fig2_380396980
- [www16] <https://arxiv.org/pdf/2406.02622>
- [www17] <https://www.ibm.com/think/tutorials/llm-safeguards-granite-guardian-risk-detection>
- [www18] <https://www.confident-ai.com/blog/the-comprehensive-llm-safety-guide-navigate-ai-regulations-and-best-practices-for-llm-safety>
- [www19] <https://www.lesswrong.com/posts/HxRjHq3QG8vcYy4yy/the-stochastic-parrot-hypothesis-is-debatable-for-the-last>
- [www20] <https://medium.com/@freddyayala/llms-are-not-stochastic-parrots-how-large-language-models-actually-work-16c000588b70>
- [www21] https://www.linkedin.com/posts/beam-data_gpt5-ai-llm-activity-7360671286517264385-FQon
- [www22] <https://uxplanet.org/3-essential-techniques-that-will-make-the-most-of-chatgpt-5-40f7141fb3bf>
- [www23] <https://aiseekhegaindia.com/docs/language-models/llmjudgement/>
- [www24] <https://tiktokenizer.vercel.app/>
- [www25] <https://www.youtube.com/watch?v=zduSFxRajkE>
- [www26] <https://www.bbc.com/news/technology-58707641>
- [www27] https://www.cnmv.es/DocPortal/Publicaciones/MONOGRAFIAS/DT_96_Large_Language_Modelsen.pdf
- [www28] <https://www.science.org/doi/10.1126/sciadv.adt3813>
- [www29] https://en.wikipedia.org/wiki/ELIZA_effect
- [www30] https://en.wikipedia.org/wiki/Dr._Sbaitso
- [www31] https://en.wikipedia.org/wiki/The_Anarchist_Cookbook
- [www32] <https://www.xataka.com/robotica-e-ia/uber-haya-ha-gastado-cuatro-meses-su-presupuesto-anual-para-ia-porque-ia-nos-esta-convirtiendoadictos-a-ella>

Referencias

- [www33] <https://www.xataka.com/robotica-e-ia/tokenmaxing-tiene-su-primer-absurdo-algunos-empleados-amazon-usan-ia-solo-para-inflar-sus-metricas-tokens>
- [www34] <https://huggingface.co/datasets/HuggingFaceFW/fineweb>
- [www35] https://en.wikipedia.org/wiki/Token_maxing
- [www36] <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>
- [www37] <https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>
- [www38] <https://www.consilium.europa.eu/en/policies/artificial-intelligence-act/>
- [www39] <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>
- [www40] <https://blog.httpcs.com/en/the-advent-of-the-ai-act/>
- [www41] http://en.npc.gov.cn.cdurl.cn/2021-12/29/c_694559.htm
- [www42] https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_3_2018_territorial_scope_after_public_consultation_en.pdf
- [www43] <https://worldpopulationreview.com/country-rankings/average-height-by-country>
- [www44] <https://keydifferences.com/difference-between-forecasting-and-prediction.html>
- [www45] <https://www.iese.edu/standout/es/rol-chief-ai-officer-caio/>
- [www46] https://en.wikipedia.org/wiki/Chief_AI_officer
- [www47] <https://www.ibm.com/think/topics/chief-ai-officer>
- [www48] <https://www.lavanguardia.com/neo/20250613/10778594/alvaro-luzon-experto-inteligencia-artificial-educados-ia-esta-perjudicando-debemos-hablarle-mal-lenguaje-violento.html>
- [www49] <https://www.lavanguardia.com/neo/20250501/10630007/alan-daiich-emprendedor-experto-inteligencia-artificial-esta-comprobado-amable-chatgpt-te-da-respuestas-mas-largas-mas-precisas-pv/vl.html>
- [www50] <https://www.amazon.es/Snake-Oil-Artificial-Intelligence-Difference/dp/069124913X>
- [www51] <https://www.youtube.com/watch?v=ieYadopst4s>
- [www52] <https://www.youtube.com/live/pmsTzFLM8rl>
- [www53] <https://www.aimeecoza.com/human-made/>
- [www54] https://en.wikipedia.org/wiki/Al_slop
- [www55] <https://www.aepd.es/preguntas-frecuentes/4-dpd/1-delegado-de-proteccion-de-datos>
- [www56] https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/roles-and-functions-of-ethics-advisory-ethics-advisory-boards-in-ec-funded-projects_he_en.pdf
- [www57] <https://www.technologyreview.com/2023/02/08/1068068/chatgpt-is-everywhere-heres-where-it-came-from/>
- [www58] <https://papers.nips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>
- [www59] <https://aighost.co.uk/pac-man-ghost-ai-how-the-classic-games-enemies-think/>
- [www60] <https://www.instagram.com/reel/DYpqo7axR3Q/>
- [www61] <https://firebrand.training/en/blog/interesting-read-alan-turing-the-experiment-that-shaped-artificial-intelligence>
- [www62] <https://medium.com/@piyushkashyap045/safeguarding-large-language-models-a-comprehensive-guide-to-enhancing-trustworthy-ai-21628ae4bf19>
- [www63] <https://www.ibm.com/think/tutorials/llm-safeguards-granite-guardian-risk-detection>
- [www64] https://www.linkedin.com/posts/theinformation_meta-employees-may-have-embraced-the-idea-activity-7449471783906484225-87UP