



<https://stablediffusion.com/>

"Create an image of Generative Artificial Intelligence like yourself in which you are like a humanoid robot, using science fiction comic style"

Introducció a la IA Generativa

UB,
7 Febrer 2025

PhD. Eloi Puertas i Prats. Facultat de Matemàtiques i Informàtica. Universitat de Barcelona

Què és la Intel·ligència Artificial

- **Intel·ligència Artificial** (1956-): La teoria i desenvolupament de **sistemes informàtics** capaços de fer tasques les quals típicament requeririen d'intel·ligència humana. (Oxford Dictionary) -> **Automatització de processos complexos**

1956 Dartmouth Conference: The Founding Fathers of AI



John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More



Alan Turing (1912-1954)

“Computing Machinery and Intelligence”

Mind - October 1950

Què és la Intel·ligència Artificial

- **Intel·ligència Artificial** (1956-): La teoria i desenvolupament de **sistemes informàtics** capaços de fer tasques les quals típicament requeririen d'intel·ligència humana. (Oxford Dictionary) -> **Automatització de processos complexos**
- **Informàtica** (1957-): Tractament automàtic de la informació (**Informació** + Automàtica) utilitzant ordinadors. -> **Automatització de processos repetitius.**

Què és la Intel·ligència Artificial

- Camp de recerca dins de **l'Enginyeria Informàtica / Computer Science**

Àrees	Problemes
Aprenentatge Automàtic (Machine Learning)	Classificació automàtica
Processament del Llenguatge Natural	Reconeixement del llenguatge
Planificació i presa de decisions	Vehicles autònoms
Visió per Computador	Reconeixement facial
Robòtica	Robots humanoides

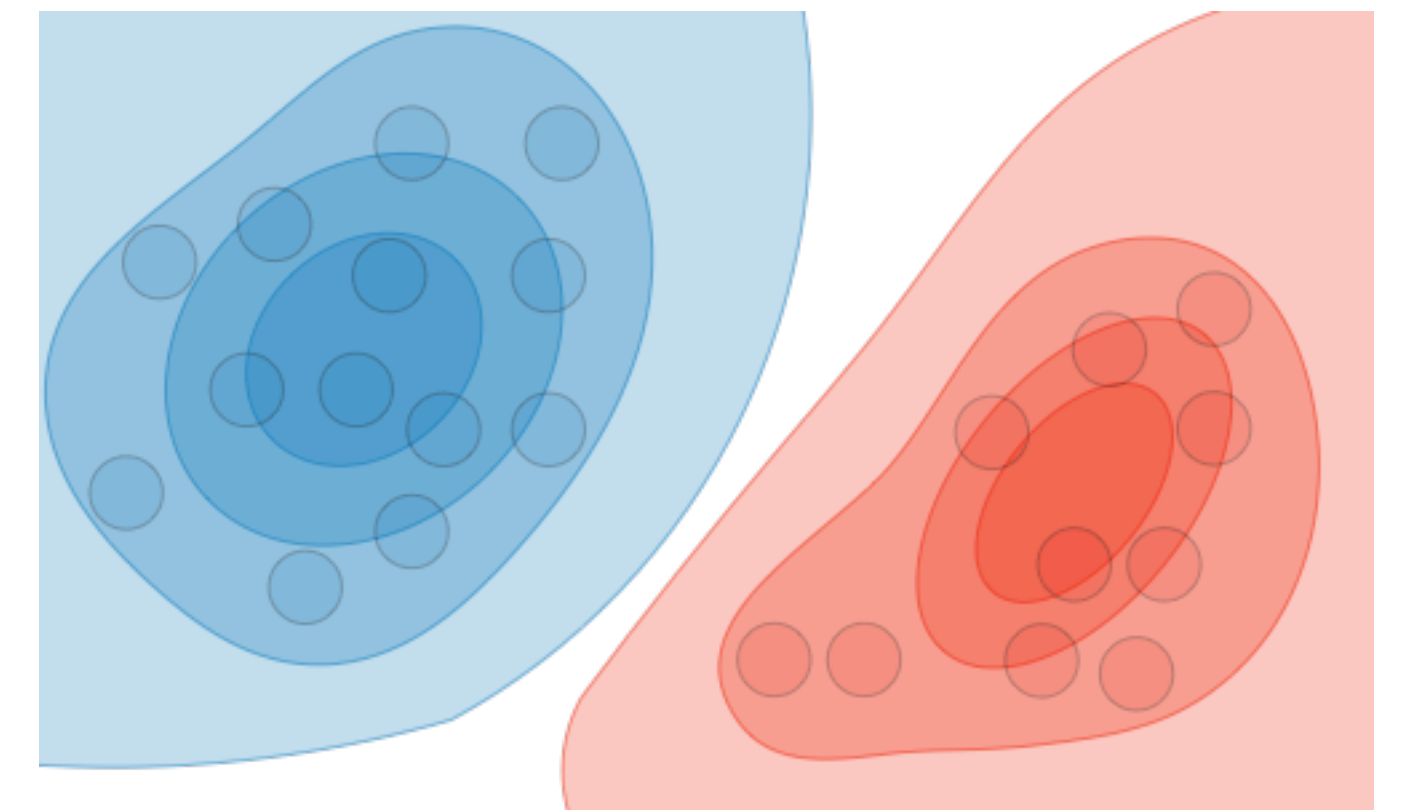
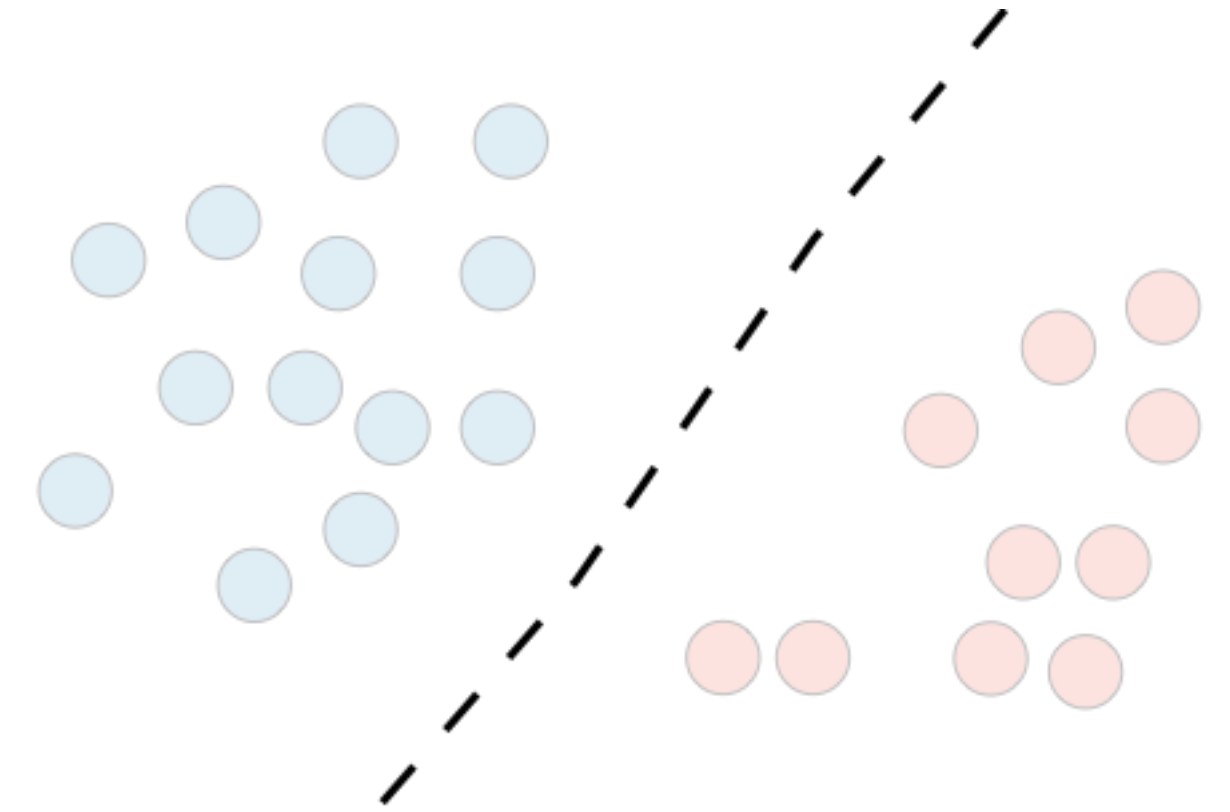
Models de la Intel·ligència Artificial

Any	Models Destacats	Tecnologia clau
1950	Cerca i optimització	Computadores
1980	Classificadors i mètodes d'aprenentatge estadístic	PCs
1990	Mètodes probabilístics	WWW
2010	Aprenentatge profund	GPUs
2020	Transformers (GPT)	Computació en el núvol.

+DADES +PROCESSAMENT = Algorismes més potents

Tipologia de problemes de la IA

- **Problemes Discriminatius:** A partir de les **dades** en fem una **classificació**. N'aprenem com distingir classes.
 - **Models supervisats**, les dades estan etiquetades
- **Problemes Generatius:** A partir de les **dades** generem **nou contingut**. N'aprenem la distribució de les dades per crear-ne nous exemples
 - **Models estadístics no supervisats**, les dades no tenen perquè estar etiquetades.



Models de Llenguatges Massius (LLM)

- Large Language Model (LLM)
 - **Tipus de problema:** Model Generatiu.
 - **Camp:** Processament del Llenguatge Natural
 - **Models usats:** Xarxes Neuronals basades en Arquitectura Transformer.
 - **Tecnologia principal:** Clústers GPUs (Dades Massives)
 - **Tasca:** A partir d'un text d'entrada "*prompt*", generar les paraules que segueixen.
 - **Aplicacions:** Bots conversacionals, Assistents, Interfície home/màquina, Processament del llenguatge natural.
- Exemples
 - **GPT** (2,3,4) OpenAI (Microsoft)
 - **Gemini** (Google)
 - **Grok** (xAI)
 - **LLAMA** (Meta)
 - **Claude** (Anthropic)
 - **Mistral** (Mistral AI)
 - **DeepSeek-R1** (DeepSeek)
 -

GPT com funciona

GPT: **G**enerative **P**re-trained **T**ransformer.

- **TASCA:** Predir quina serà la següent paraula més probable i afegir-la a l'entrada:

"la suma dels quadrats dels catets és igual a"
→ "la "

"la suma dels quadrats dels catets és igual a la"
→ "suma" → ~~"dels quadrats dels altres dos costats"~~
→ "hipotenusa" → "al quadrat"
→ "longitud" → "del costat hipotenusa al quadrat"

prompt GPT3.5: Completa les tres frases més probables després de la frase: "la suma dels quadrats dels catets és igual a"

- Usa model preentrenat amb **grans quantitats de dades** no etiquetades: Llibres, articles, documents obtinguts d'internet...
- Nombre de **paràmetres** del model LLM (mida del model, nombre de neurones de la xarxa neuronal)

Model	Paràmetres
GPT-1	117 Milions
GPT-2	1.5 Bilions
GPT-3	175 Bilions
GPT-4	1.76 Trilions

- A més a més: Aprenentatge per reforç mitjançant **retroalimentació humana**

Limitacions. Què i què no pot fer un LLM

Tasques que es poden fer amb LLM

- **Generació i modificació de textos**
- **Traducció automàtica.**
- **Resumir documents.**
- **Generar preguntes i respostes sobre textos.**
- **Mantenir una conversa.**
- **Respondre a preguntes.**
- **Tasques típiques de Processament de Llenguatge Natural**
 - **Reconeixement d'entitats**
 - **Anàlisi del sentiment**
 - **Classificació de textos.**
- **Generar imatges, vídeos, música, codi...**

Limitacions. Què i què no pot fer un LLM

Limitacions Principals dels LLM simples:

- No disposa de **sentit comú ni d'enteniment**
 - **Lloro** estocàstic (Stochastic Parrot )
 - No pot entendre **contextos implícits** o “llegir entre línies”
 - No te capacitat de **raonament** o fer càlculs per si sol
- **Per si sol, no te accés a dades** fora del model
 - **Coneixement limitat** als documents que te el model (**generalista**)
 - Respostes **esbiaixades**
 - Falta de **profunditat i creativitat**
- Limitacions **tècniques**:
 - No pot fer més d'**una tasca a l'hora**
 - No pot fer tasques **que no generin** res.
 - Pot donar respostes **incompletes** o **al·lucinar**
 - Es sol necessitar fer un ***fine-tuning*** per aconseguir millor resultats
 - **Consumeix** gran quantitat de recursos, 30x que una consulta a google.

Beneficis i Riscos dels models LLM

- Beneficis:

- Augmentar productivitat en tasques repetitives.
- Facilitat d'ús.
- Poder-hi interactuar per millorar els resultats.
- Usar-lo com a assistent o company.
- Models bàsics gratuïts.

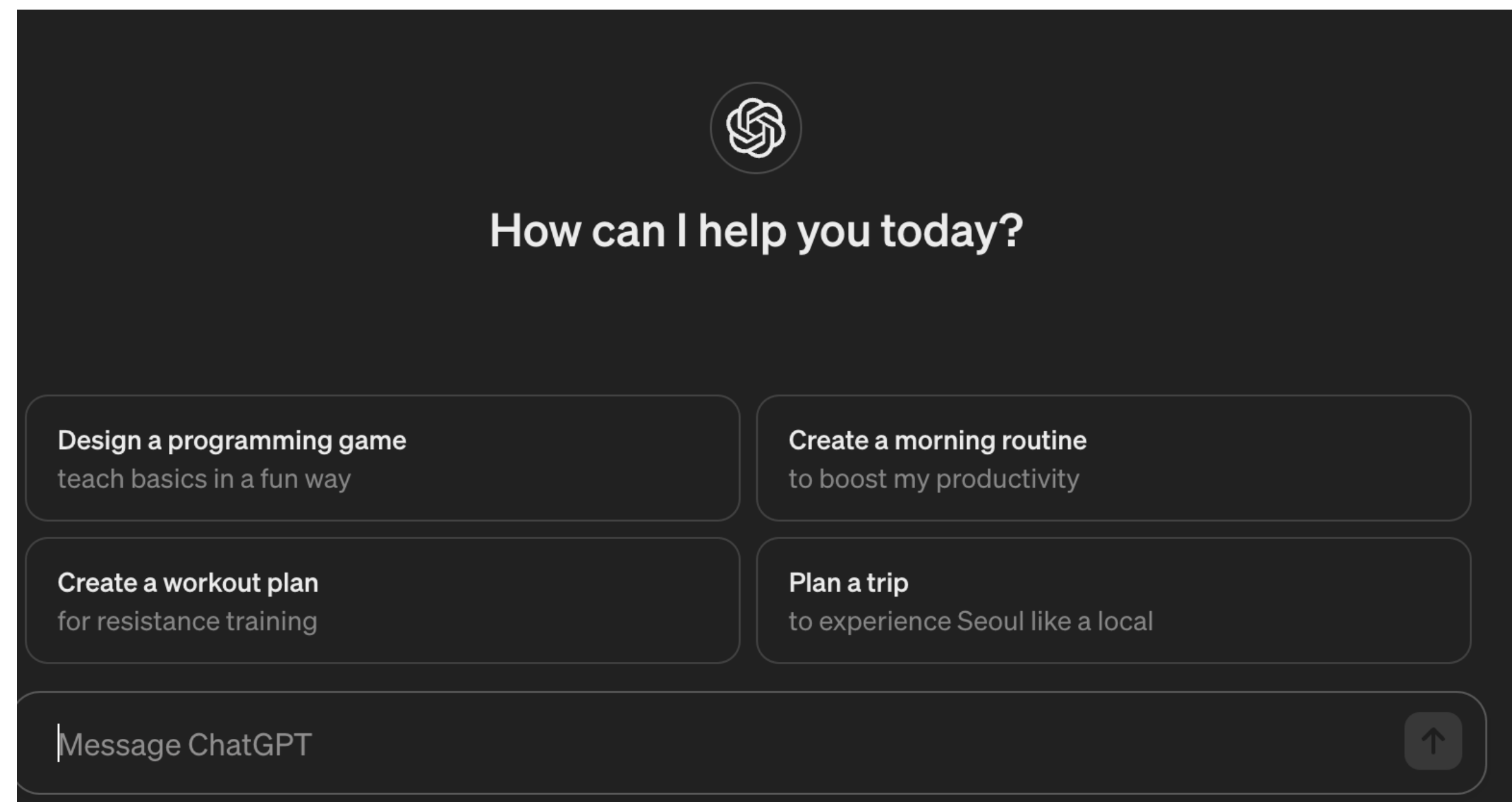
- Riscos:

- Resultats poden ser no acurats, ambigus o fins i tot amb errors.
- Usar-lo de forma negligent.
- Plagiar els resultats obtinguts pel GPT
- Propietat intel·lectual de les dades usades i dels resultats obtinguts

ChatGPT: Prompt

- **ChatGPT: Aplicació WEB**

d'Intel·ligència Artificial generativa que usa el model GPT d'OpenAI



- **“PROMT”**: Seqüència de text o qualsevol altre mitjà que s'utilitza com a instrucció per a l'entrada del model GPT

- Elements bàsics:

Context - informació externa o context addicional que pot dirigir el model cap a respostes més adequades

Instrucció - una tasca específica o instrucció que vols que el model realitzi

Dades d'entrada -dades addicionals per tal d'ajudar al model a trobar una resposta

Indicador de sortida - el tipus o format de la sortida.

Prompt Engineering

- “Desenvolupament i optimització de *prompts* per utilitzar eficientment els models de llenguatge massius (LLM) en una àmplia varietat d'aplicacions i temes de recerca.”
- Guies d'enginyeria de prompts:
 - <https://www.promptingguide.ai/ca>
 - <https://platform.openai.com/docs/guides/prompt-engineering>

Consells bàsics

- Començar amb **coses senzilles**
- **Especificar clarament** la **tasca** que es vol realitzar: “Escriu”, “Classifica”, “Resumeix”, “Tradueix”, “Ordena”...
- Ser el **màxim descriptiu** possible
- Evitar **imprecisions**

Consells bàsics



You

###Context###

Ets un agent traductor britànic

Instrucció

Tradueix el següent text a l'anglès:

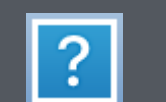
Text: "hola! com va?"



ChatGPT

Text:

"Hello! How are you?"



You

Extreu el nom de llocs en el següent text

Format desitjat:###

"<Llista separada per comes de llocs>"

Entrada:

"Ens trobarem a les 17:00 davant del bar Zurich de Plaça Catalunya per anar després al FNAC"



ChatGPT

Llocs: bar Zurich,
Plaça Catalunya,
FNAC

Consells bàsics



You

Explica el concepte física quàntica. Mantingueu *l'explicació breu, només unes quantes frases, i no siguis massa descriptiu.*



You

Utilitza **2 frases** per explicar el concepte física quàntica a un **estudiant de secundària.**



ChatGPT

La física quàntica és el camp de la física que estudia el comportament dels objectes a l'escala microscòpica, com els àtoms i les partícules subatòmiques. S'utilitza per comprendre fenòmens com l'entrellaçament i la superposició, que són diferents de la física clàssica.



ChatGPT

La física quàntica és una branca de la física que explora el comportament dels objectes molt petits, com els àtoms i les partícules subatòmiques. Desafia les nostres intuïcions, introduint conceptes com l'entrellaçament i la superposició, que són diferents del que esperaríem segons la física clàssica.

Prompts avançats

- *Few-shot prompting*: proporcionar exemples de com o què hauria de respondre



You

###Instrucció###

Quin és el sentiment d'aquestes frases:

Dades entrada

Això és genial! -> Positiu

Això és dolent! -> Negatiu

Vaja, quina pel·lícula més xula! -> Positiu

Text

Quin espectacle més horrible! -> ?



ChatGPT

El sentiment de la frase "Quin espectacle més horrible!" és negatiu.

Prompts avançats

- *Cadena de pensament (CoT)*: Permet habilitats de raonament més complex a través d'indicar-li passos intermitjos



You

Digues com està la següent ciutat de lluny de Terrassa.

Primer: calcula la distància en km en que es troba Terrassa de les ciutats de la llista d'exemples i mostra-ho.

Segon: Associa els km amb com de lluny està etiquetada en la llista d'exemples.

Tercer: Digues com de lluny està la ciutat que falta per etiquetar

Quart: Assigna la millor etiqueta de les dels exemples

Barcelona: A prop

Manresa: A prop

Sabadell: Al costat

Lleida: Lluny

Madrid: Molt Lluny

Girona: ?



ChatGPT

Primer: Distància entre Terrassa i les altres ciutats:

Barcelona: 28 km

Manresa: 50 km

Sabadell: 9 km

Lleida: 140 km

Madrid: 625 km

Girona: 107 km

Segon: Associació dels km amb la proximitat:

Barcelona: A prop (28 km)

Manresa: A prop (50 km)

Sabadell: Al costat (9 km)

Lleida: Lluny (140 km)

Madrid: Molt Lluny (625 km)

Girona: Lluny (107 km)

Tercer: La ciutat de Girona està a 107 km de Terrassa.

Quart: Girona es podria etiquetar com "Lluny".

Prompts avançats

- *Cadena de pensament (CoT)*: Permet habilitats de raonament més complex a través d'indicar-li passos intermitents



You

Pas 1. Vaig anar al mercat i vaig comprar 10 pomes.

Pas 2. Vaig donar 2 pomes al veí i 2 al meu germà.

Pas 3. Després vaig anar a comprar 5 pomes més i em vaig menjar 1.

Amb quantes pomes em vaig quedar?

Fes-ho pas a pas.



ChatGPT

Pas 1: Inicialment, vaig tenir 10 pomes.

Pas 2: Vaig donar 2 pomes al veí i 2 al meu germà.

Això deixa $10 - 2 - 2 = 6$ pomes.

Pas 3: Després de comprar 5 pomes més, tindrè $6 + 5 = 11$ pomes. Després de menjar-me 1, em queden $11 - 1 = 10$ pomes.

Per tant, em vaig quedar amb 10 pomes.

Prompts avançats

Arbre de pensament (ToT) Prompting. Fer actuar diferents “agents” pas a pas.

Nous models com **GPT o1** ho apliquen internament



You

Actua com a tres experts diferents que executen una tasca que se'ls encomana

Pas 1. Tots els experts anotaran la seva resposta.

Pas 2. Compartiran amb el grup la justificació

Pas 3. Dona una resposta final consensuada

La tasca és digues si aquesta afirmació es certa o no: La terra és plana



ChatGPT

Expert 1: Geògraf

Pas 1: La terra no és plana.

Pas 2: Justificació: Com a geògraf, sé que la Terra és un cos esfèric. Hi ha moltes proves científiques, com ara l'observació dels moviments estel·lars, els viatges espacials i les observacions satel·litàries, que confirmen la forma esfèrica del nostre planeta.

Expert 2: Astrònom

Pas 1: La terra no és plana.

Pas 2: Justificació: L'astronomia ens mostra que la Terra és un cos esfèric. A través de l'observació de les fases de la lluna, els eclipsis, i altres fenòmens celestials, podem veure que la Terra projecta una ombra circular quan està il·luminada pel sol, cosa que indica clarament la seva forma esfèrica.

Expert 3: Físic

Pas 1: La terra no és plana.

Pas 2: Justificació: Des d'un punt de vista físic, les lleis de la gravetat i la mecànica celeste expliquen la forma esfèrica de la Terra. La distribució de la massa i l'atracció gravitatòria condueixen a la formació d'un cos esfèric. A més, les observacions de la corba de l'horitzó i els fenòmens com els viatges circumnavegants han demostrat que la Terra és arrodonida.

Resposta Final Consensuada: La terra no és plana.

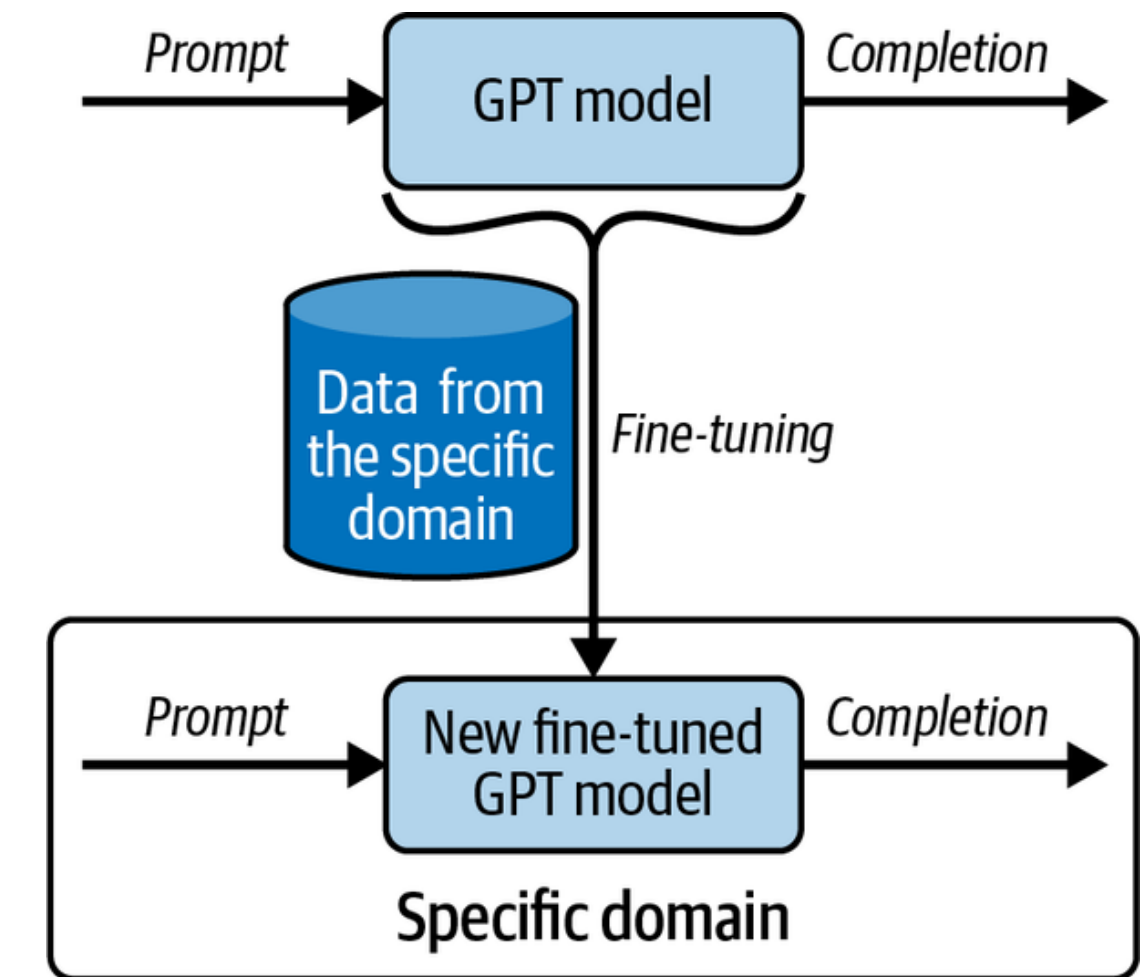
Totes les proves i raonaments proporcionats pels experts coincideixen en afirmar que la Terra és un cos esfèric.

Ajustar els LLMS

- **Ajustar** el model de LLM per a tasques específiques introduint **dades pròpies**.
 - Es poden usar **més dades per entrenar** que les que caben en un *few shot prompt*.
 - Es poden fer aplicacions dedicades: **Bots, Assistents, GPTs...**
 - Tasques que requereixen d'ajustos:
 - Anàlisi de dades de recerca
 - Creació de chatbots de dominis específics
 - Recuperació d'informació de dades pròpies.
- **Opció 1:** *Fine Tuning* del model
 - **Opció2:** *Retrieval Augmented Generation*

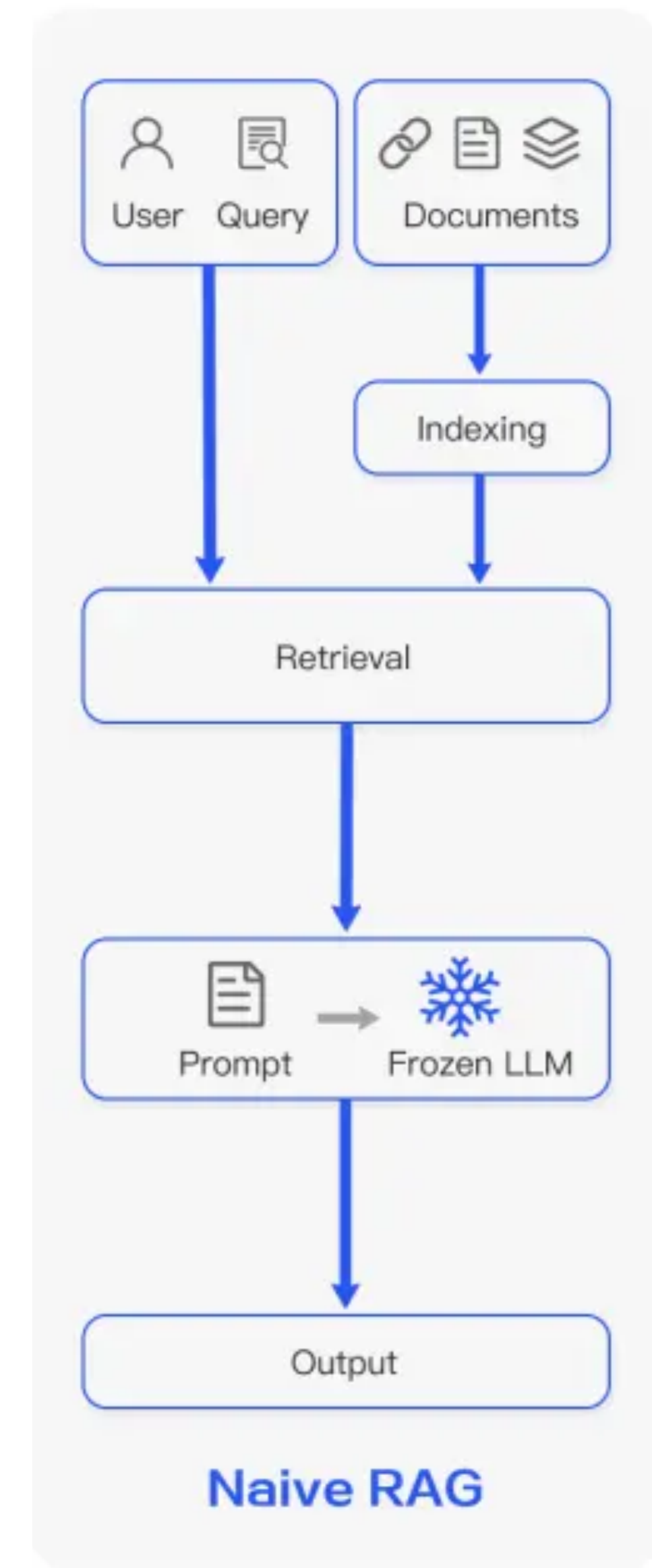
Fine Tuning

- **Afinar** el **model de LLM** per a tasques específiques introduint **dades pròpies**.
- Es **redueix la longitud** dels prompts ja que el model estarà afinat al nostre problema
- **És costós** en temps i energia, sol tenir un gran cost econòmic.



RAG (Retrieval Augmented Generation)

1. L'usuari **envia** el *prompt*
2. Es fa una **cerca** d'informació rellevant sobre el domini del *prompt* en una **base de dades externa**
3. **S'enriqueix** el *prompt* amb les **dades extretes** de la Base de Dades.



PlayGrounds

- Plataformes online per a provar diferents models i prompts, més enllà de l'aplicació chatGPT
 - OpenAi Playground: Ideal per provar *prompts* complexos amb chatGPT, tens més opcions.
 - Vercel: Ideal per comparar el mateix *prompt* en diferents models
 - Poe: Ideal per experimentar amb diferents bots. N'hi ha d'oficials i d'altres d'especialitzats. Pots crear el teu propi assistent
 - HuggingChat: Si es vol tenir una alternativa a ChatGPT que utilitzi models d'ús lliure.
 - Chatbot Arena: Per veure diferents classificacions entre els LLMs més usats.
 - FlowGPT: Per crear els teus propis bots i interactuar amb moltíssims creats pels usuaris.

LLM multimodals

- Es tracta de LLM que no només reben com a entrada text i generen text, sinó que poden rebre i generar diferents tipus de mèdia.
 - **GPT4** és un model LLM multimodal que pot rebre text, imatge com a prompt i pot generar imatges i text.
 - **Gemini** (Google) i **SeamlessM4t** (Meta) també són Multimodals
- Altres LLMs estan especialitzats en dominis específics, més enllà de Text-to-Text:
 - Àudio-Text: **Whisper** (OpenAI)
 - Text-Image: **Dall-e** (OpenAI), **Stable Diffusion** (Stability AI), **Midjourney** (Midjourney)
 - Text-Codi: **Codex** (OpenAI/Microsoft)
 - Text-Video: **Syntheisa**, **Lumen5**, **Sora** (Open AI)
 - Text-Àudio: **Suno**, **Stable Audio** (Stability AI)

Detecció de contingut generat per AI

Existeixen detectors d'IA?

- AI Text Classifier (OpenAI): *As of July 20, 2023, the AI classifier is no longer available due to its low rate of accuracy.* 😱
- *Educator FAQ OpenAI: Do AI detectors work?*
 - *In short, no, not in our experience.*
 - Detecta molts falsos positius, textos humans etiquetats com generats per l'AI, per exemple, un fragment de Shakespeare que hagi estat usat per al seu entrenament.

Detectors d'IA Generativa

Existeixen però productes que podeu experimentar per si voleu confirmar sospites:

- [CopyLeaks](#)
- [Content at Scale](#)
- [GPTZero](#)

El problema és que ja existeixen aplicacions d'IA generativa que Humanitzen els textos generats:

- [Humanize.ai](#)

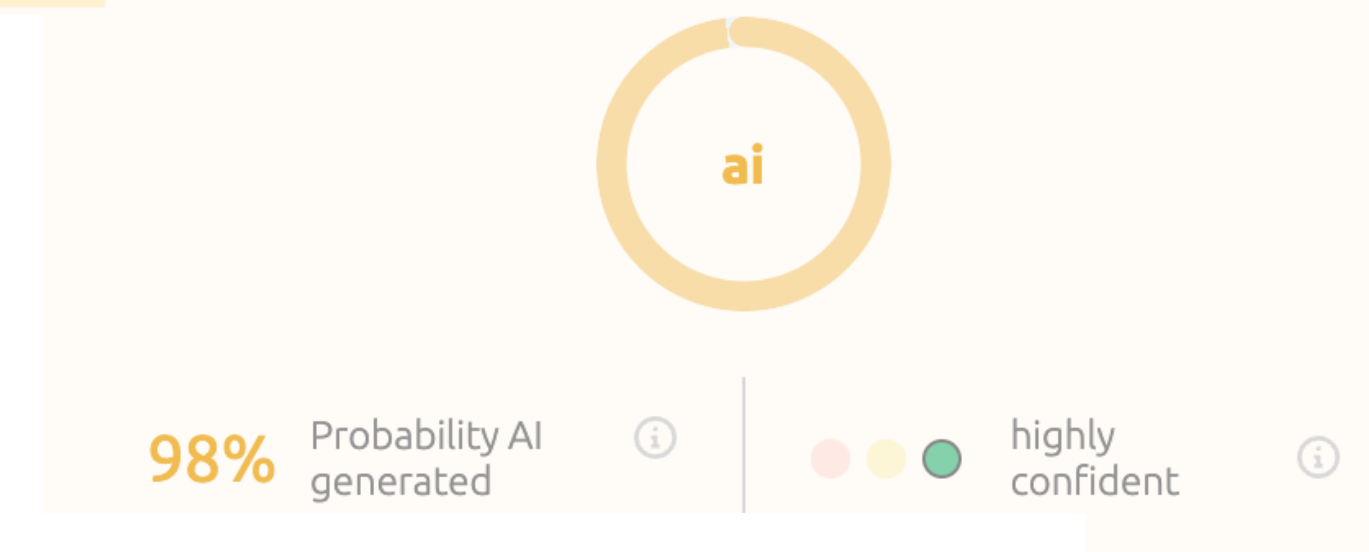
ChatGPT 3.5

Sara's dedication extended beyond her coursework, as evidenced by her successful getting a grant for a collaboration in my research group in Machine Learning and Data Science. During her stay, she applied cutting-edge methodologies to analyze temporal educational data, focusing on patterns in student and subject performance over the years, including the unique challenges posed by the COVID-19 pandemic.

Humanize

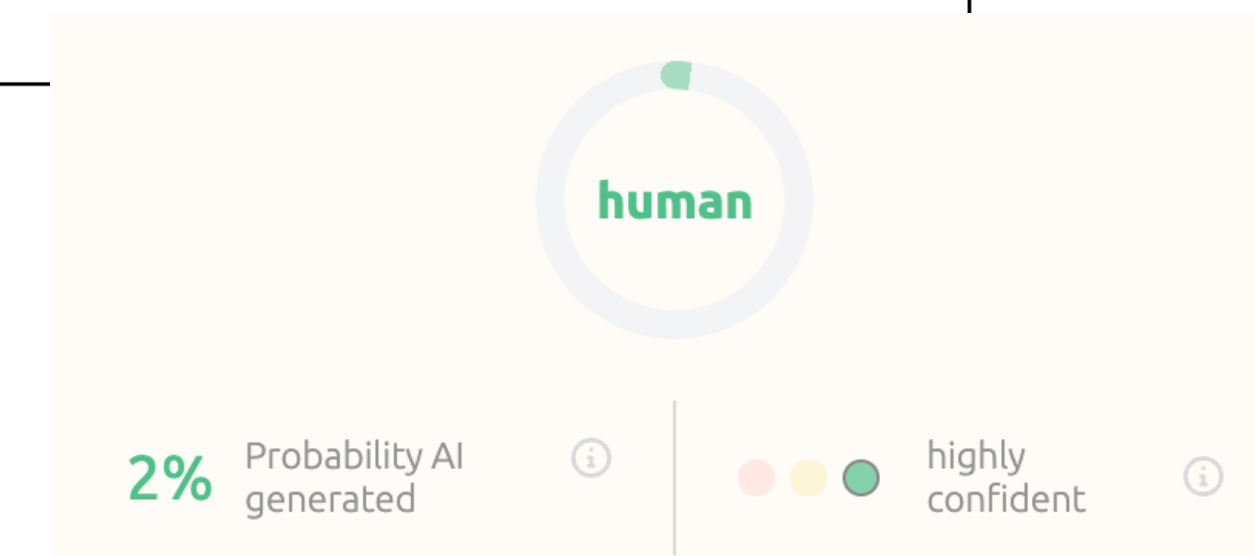


GPTZero



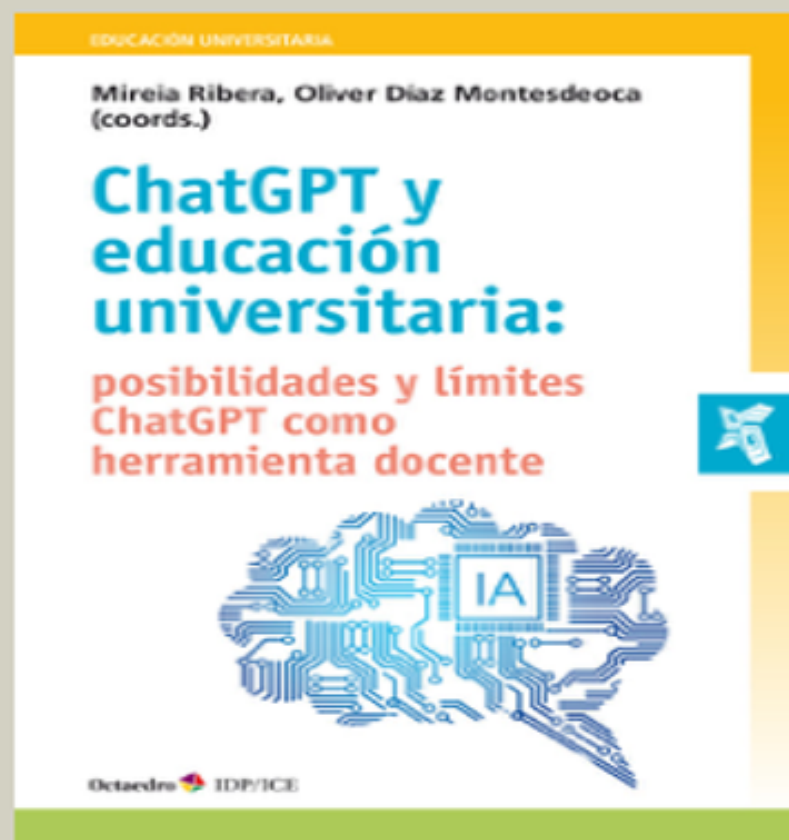
In addition to her academic work, Sara's commitment was demonstrated by the fact that she won a grant for a project she collaborated on for my research group in Machine Learning and Data Science. While she was here, she got to apply state-of-the-art techniques in analysing temporal educational data looking at trends on student performance and subject performance for several years including challenges posed by COVID-19.

GPTZero



Per saber-ne més...

Octaedro
Editorial



Colección Educación Universitaria

Código PDF: 1-15224

OPEN ACCESS 

ISBN: 978-84-10054-01-1

Para más información

93 246 40 02

octaedro@octaedro.com

www.octaedro.com

ChatGPT y educación universitaria

Posibilidades y límites ChatGPT como herramienta docente

Mireia Ribera y Oliver Díaz Montesdeoca (coords.)

Casi todos hemos jugado ya con ChatGPT, pero ¿le sacamos el máximo provecho? ¿Nos hemos parado a pensar sobre sus implicaciones en la docencia? ¿Cuáles son las limitaciones y retos éticos de esta herramienta? ¿Cómo podemos mejorar los *prompts*?

ChatGPT es mucho más que una herramienta de pregunta-respuesta. Es un cambio disruptivo y acelerado en los mecanismos y estrategias de búsqueda de información, comunicación y de creación de contenido docente. Es un desafío perturbador, que, en estos momentos todavía iniciales, genera enormes dudas entre el profesorado, pero a la vez, supone una oportunidad única para abrir nuevos caminos de aprendizaje. ¿Te animas a explorar sus posibilidades?

Un libro que pretende acompañaros en este camino y que no puede faltar en la biblioteca del educador inquieto. Una visión rica y multidisciplinar, que sienta las bases para entender cómo funciona, hacia donde puede evolucionar esta tecnología y nos acompaña en la reflexión de su uso en los procesos de enseñanza-aprendizaje. Con el compromiso de actualizaciones regulares según evolucione la tecnología y maduren las estrategias de su implementación.

COORDINACIÓN

Mireia Ribera es doctora en Documentación Digital y profesora del Departamento de Matemáticas e Informática de la Universidad de Barcelona (UB). Su formación interdisciplinaria le permite tener una visión crítica, aunque constructiva, de las innovaciones tecnológicas y analizar sus implicaciones en la sociedad.

Oliver Díaz Montesdeoca es doctor en Ingeniería electrónica por la Universidad de Surrey y profesor del Departamento de Matemáticas e Informática de la UB. Su experiencia en Inteligencia artificial confiable aporta una visión tecnológica centrada en el usuario final.

AUTORÍA

Mariona Grané es investigadora del grupo LMI y profesora en la Facultad de Educación de la Universidad de Barcelona. Su campo de estudio se centra en la integración de las tecnologías digitales en entornos educativos para el aprendizaje y la construcción de conocimiento.

Daniel Ortiz Martínez es doctor en Inteligencia artificial por la Universidad Politécnica de Valencia. Es profesor del Departamento de Matemáticas e Informática de la UB. Gran conocedor del procesamiento del lenguaje aporta una visión técnica de ChatGPT.

David Buchaca es doctor en Inteligencia Artificial por la UPC y trabaja actualmente en Amazon, tras una prolífera carrera en Jina, Barcelona Super Computing Center e IBM. Conocedor de la aplicación de la IA en la industria, aporta una visión aplicada y realista de los retos y posibilidades de esta.

Eloi Puertas i Prats es doctor en Inteligencia artificial por la UB. Es profesor del Departamento de Matemáticas e Informática de la UB. Sus intereses en la ciencia de datos aplicada al aprendizaje son claves para explicar el uso de esta herramienta.

Carlos Lopezosa es doctor en periodismo por la Universitat Pompeu Fabra e investigador visitante en la UB. Su interés en los medios de comunicación online ayuda a entender los posibles usos de esta tecnología y sus limitaciones.

Javier Guallar es doctor en Información y Documentación por la UB y doctor en Comunicación por la Universidad Pompeu Fabra. Es profesor de la Facultad de Información y Medios Audiovisuales de la UB. Su experiencia en la curación de contenidos, periodismo digital y la documentación periodística aporta una visión global sobre el ChatGPT.



2. Indicaciones prácticas para usar ChatGPT (*Eloi Puertas i Prats*)

3. La tecnología tras ChatGPT (*Daniel Ortiz, David Buchaca i Prats*)