



# AMB L'ARRIBADA DE LA IA GENERATIVA EL "PROMPT ENGINEERING" S'HA CONVERTIT EN LA NOVA COMPETÈNCIA EDTECH A ADQUIRIR. CONSIDEREU QUE EL PROFESSORAT I ALUMNES TENEN PROUS INDICACIONS SOBRE ELS SEUS ABUSOS I MALS USOS (PER EX. CONTINGUT MALICIÓS)?

- La gran majoria (85%) considereu que el professorat i alumnes no tenen prou indicacions sobre els abusos i mals usos de la IA generativa i el "prompt engineering", mentre que la resta no ho sabeu (7%) o sou ambivalents (5%).
- Així doncs, els "enginyers ràpids" són persones que escriuen prosa més que no pas codi per posar a prova els resultats dels sistemes d'IA generativa, i per detectar que les seves respostes siguin reproduïbles i que segueixin protocols de seguretat.
- Per tant, si bé aquest perfil de professional respon a la necessitat d'avaluar els models d'interacció home-màquina, amb l'arribada del Chat GPT es pot dir que tothom s'ha convertit en una espècie de "enginyer ràpid", però no pas per avaluar els sistemes d'IA generativa.
- De fet, més que avaluar el que estem fent és entrenant-los amb preguntes cada vegada més específiques, personals i críptiques, un fet que comporta riscos com la generació de contingut esbiaixat, de desinformació i també, és clar, de contingut maliciós.
- Davant d'aquesta situació, ja podem intuir un doble problema que cal abordar. Per una banda, calen més persones que facin la veritable feina d'avaluar els models d'interacció home-màquina. Per una altra banda, calen més indicacions sobre els seus abusos i mals usos.
- Això és especialment rellevant en àmbits com el de la salut o l'educació, on els consells i continguts personalitzats sempre apareixen com més atractius per un creixement i aprenentatge individualitzat de les persones. Però al marge d'aquests beneficis també hi ha inconvenients.





- La principal preocupació, especialment en l'àmbit de l'educació, és el potencial de biaix en el contingut educatiu que es genera, entre altres coses perquè els algorismes només són tan imparcials com les dades en què s'entrenen. I, recordem, que la IA generativa va d'entrenament.
- Per exemple, davant les evidències ja no podem descartar una IA generativa creada per determinats grups segons l'ètnia, gènere o estat socioeconòmic, fent que es reforcin estereotips negatius i afectar negativament l'experiència d'aprenentatge.
- D'altra banda, com els sistemes d'IA generativa estan limitats per les dades que se'ls han proporcionat, això significa que no poden ni tenen la capacitat de produir contingut innovador i a la vegada crític en les respostes, malgrat una publicitat que ho postula per raons comercials.
- Per tant, el potencial de confiar massa en la IA generativa és evident, ja que proporciona respostes i solucions aparentment immediates a l'aprenentatge. No obstant, els reptes de validació de fonts, de continguts i de si es pot incórrer en possible plagi romanen intactes.
- A aquests problemes cal afegir el de la bretxa digital entre els estudiants i professorat amb i sense accés a la tecnologia d'IA generativa, la qual cosa pot augmentar la bretxa d'aprenentatge. Sabem que els mals usos de la IA generativa dependran en gran mesura de les capacitats.
- Així doncs, sense abordar-los a nivell individual i col·lectiu, no tan sols hi haurà problemes de fiabilitat i credibilitat, de validesa i veracitat i de plagi i ús honest de la informació, també hi ha el risc de generar noves divisions entre estudiants i professorat.
- Per això, hem de recordar que els educadors continuen jugant un paper insubstituïble en el procés d'aprenentatge. Si bé la IA generativa pot tenir un rol en l'educació, no tindrà ni la creativitat ni l'originalitat ni el pensament crític que les persones poden oferir.
- Per tant, si cal aprofitar la IA generativa ha de ser per beneficiar el desenvolupament d'habilitats de l'aprenentatge de tothom i garantir que, quan aquesta s'introdueix a l'aula, no dificulta el desenvolupament del pensament crític i les habilitats de resolució de problemes.







- Entre les qüestions de pensament crític que cal emfatitzar, és si cal utilitzar la IA generativa de manera constant, ja que això significa una ingent quantitat diària d'emissions de CO2 que prioritza les nostres experiències immediates sobre les conseqüències a llarg termini.
- I fins aquí per avui. Us recomanem escoltar la [sèrie de podcasts de Education International sobre l'educació en un món on les grans empreses #EdTech dirigeixen les agendes educatives.](#)
- I, és clar, [el nostre circuit OEIAC sobre l'ús responsable i sostenible del #ChatGPT.](#)

**Què és el ChatGPT i quines consideracions ètiques hem de tenir?**

**Definició**  
ChatGPT és un prototip de bot conversacional d'intel·ligència artificial que, fonamentalment, intenta generar una "continuació raonable" de qualsevol text que s'hagi escrit fins ara, on per "raonable" entenem tot el que la gent ha escrit en milers de milions de pàgines web, etc. (Wolfram 2023).

**G-P-T**  

- Generative** → permet la generació paraula a paraula
- Pre-trained** → entrenat prèviament amb grans quantitats de dades
- Transformer** → xarxa neuronal artificial basada en un model codificador-decodificador

**Capacitats i limitacions**  
 Segons OpenAI, el ChatGPT té les següents capacitats:
 

- Recordar allò que l'usuari ha dit en el transcurs de la conversa.
- Permetre a l'usuari oferir correccions sobre el resultat.
- Declinar peticions inapropiades.

 Segons OpenAI, el ChatGPT té les següents limitacions:
 

- De vegades pot generar informació incorrecta.
- De tant en tant pot produir instruccions perjudicials o contingut ofensiu.
- Té un coneixement limitat del món i dels esdeveniments posteriors a 2021.

**Has provat el circuit OEIAC sobre l'ús responsable i sostenible del ChatGPT?**

Fonts: Wolfram 2023, Kulkarni 2023, Tsuklenov 2023, Pointon 2022, Haggart 2023, Perrigo 2023, OpenAI

