39è CICLE DE CONFERÈNCIES 2023: 37-44 (2024)
Societat Andorrana de Ciències, Andorra
DOI: 10.2436/15.8060.35.3 ISBN: 978-99920-61-74-9

La irrupció de la intel·ligència artificial a través dels Models de Llenguatge Extens com el ChatGPT

· 21 d'abril del 2023 a les 7 del vespre
· Sala d'actes del Centre Cultural La Llacuna, Andorra la Vella
· Taula rodona
· Amb el suport de Creand© Crèdit Andorrà

Jordi Herrera i Joancomartí

Doctor en Matemàtica Aplicada i Telemàtica, professor al Departament d'Enginyeria de la Informació i les Comunicacions de la UAB

Resum
El frenètic desenvolupament que, en els últims anys, estan tenint els models de llenguatge extens, popularitzats per l'aplicació del ChatGPT, han rellançat les potencialitats de la intel·ligència artificial i l'han fet arribar al gran públic. La societat, de cop, ha copsat la rellevància d'aquesta tecnologia i s'ha començat a preguntar per les implicacions que pot tenir la seva adopció. En aquest sentit, la Societat Andorrana de Ciències va organitzar l'abril del 2023 la taula rodona *Reptes de la intel·ligència artificial avui a Andorra*, de la qual l'autor d'aquest article va ser ponent. Aquest article sintetitza les aportacions que l'autor va fer en els diferents àmbits que es van debatre.

Introducció

En els últims anys, l'increment de la potència de càlcul dels ordinadors unit a l'alta disponibilitat de volums ingents de dades han permès el perfeccionament de noves eines d'Intel·ligència Artificial (IA) que utilitzen les xarxes neuronals [1] i l'aprenentatge profund [2]. Una de les eines més conegudes que s'ha posat a disposició del gran públic són els models de llenguatge extens [3] (en anglès, Large Language Models – LLM) com ara el ChatGPT [4], pioner en l'àmbit i desenvolupat per la companyia OpenAI, el Copilot de Microsoft [5] o el Gemini de Google [6], entre d'altres. Aquestes eines, basades en models estadístics, s'entrenen amb grans conjunts de dades per tal d'aconseguir models que permeten predir la sortida d'un conjunt de valors d'entrada. Aquesta breu i sintètica descripció del funcionament dels LLM és important perquè permet identificar-ne dos aspectes essencials: d'una banda, la necessitat d'un conjunt de dades d'entrenament del sistema i d'altra banda, la component probabilística del propi sistema.

Els LLM permeten processar text dels usuaris, sovint en forma de preguntes, per proporcionar-ne possibles respostes. A diferència dels cercadors convencionals, que a partir de la pregunta busquen una font de dades que contingui una temàtica coincident i la mostren, els LLM, d'alguna manera, condensen totes les possibles respostes de totes les possibles preguntes en un model probabilístic i, a partir de la pregunta, el mateix sistema elabora la resposta. I és aquesta naturalesa probabilística dels LLM la que fa que una mateixa pregunta feta dues vegades a un mateix LLM

no retorni exactament el mateix resultat. A més, el model probabilístic que s'utilitza fa que el sistema no sempre proporcioni una resposta correcta a la pregunta que estem realitzant, cosa que produeix el que en l'argot de la IA es coneix com a *al·lucinació*. Aquestes al·lucinacions en els LLM es produeixen per diferents motius, que van des del mecanisme de desenvolupament del mateix LLM fins a les particularitats del conjunt de dades que s'han utilitzat per entrenar-lo.

La importància de les dades d'entrenament

Com acabem de comentar, l'aparició dels LLM s'ha donat gràcies a la quantitat ingent de dades emmagatzemades que tenim actualment i la potència de càlcul que ens permet gestionar-les i analitzar-les. Són aquestes dades les que s'utilitzen en el conjunt d'entrenament. Això fa que, d'alguna manera, un LLM pugui condensar un conjunt immens de dades que han estat utilitzades per entrenar-lo i que utilitzarà posteriorment en les seves respostes. Ara bé, aquestes dades que aquests sistemes utilitzen en l'entrenament, al cap i a la fi, són dades generades pels humans de forma col·lectiva. Per això, hi ha qui considera que el nom d'intel·ligència artificial per referir-se als LLM no està ben triat perquè possiblement seria més acurat parlar d'intel·ligència col·lectiva. De fet, un dels cavalls de batalla comercial en les empreses d'IA actuals són justament les dades que fan possible l'entrenament d'aquests models, dades sovint generades per usuaris, i al qual més endavant ens referirem.

Deixant de banda per un moment el component probabilístic dels LLM, la qualitat de les respostes d'un LLM es basa en la qualitat de les dades amb les quals s'ha entrenat el sistema. Com més dades s'utilitzen, més bons resultats ofereix el sistema. Ara bé, el sistema només pot *aprendre* de les dades que se li proporcionen, de manera que si el conjunt de dades d'entrenament és un conjunt limitat, l'aprenentatge tindrà també un coneixement limitat. A més, com es dona en la majoria dels casos, si el conjunt de dades d'entrenament té biaixos, les respostes del sistema també seran esbiaixades en el mateix sentit. Per exemple, si entrenem un sistema amb el conjunt de sentències condemnatòries del sistema judicial dels Estats Units podem crear un sistema que associï la raça negra a un conjunt de fets delictius, pel fet que durant molts anys el biaix en les condemnes en aquest país establien aquesta relació. Aquest fet pot propiciar situacions en les quals l'ús de sistemes d'IA per a segons quines finalitats representi un problema molt greu que arribi a acusar ciutadans davant la justícia sobre la base dels resultats d'aquests algorismes [7].

La correcció de la informació proporcionada pels LLM

Com ja hem mencionat, la informació que ens proporciona un LLM prové de complexos sistemes que combinen mecanismes probabilístics amb els conjunts de dades d'entrenament utilitzats en aquests sistemes. Per tant, si les dades que entrenen aquest sistema no incorporen informació sobre la temàtica que estem tractant o el model probabilístic no funciona correctament podem obtenir respostes poc acurades o directament errònies, cosa que s'anomena *al·lucinacions*. Aquest punt té una importància crucial perquè difereix d'altres mecanismes que han anat apareixent al llarg de la història. Per exemple, quan es parla dels LLM i de la seva possible acceptació o rebuig, en ocasions es compara amb una calculadora [8] i de com aquests aparells, en la seva aparició, van presentar un rebuig per la pèrdua de capacitat de càlcul mental que suposaria per a la població. Més enllà si la població ha de tenir una certa capacitat de càlcul mental o no, el cert és que una calculadora proporciona respostes precises i, sobretot, sempre correctes, a una

operació que se li demani. A l'usuari no li cal saber com se suma ni com es resta per poder fer una d'aquestes operacions amb dos números. La resposta que obtindrà de la calculadora sempre serà correcta. La diferència amb un LLM és que la resposta obtinguda d'aquest nou sistema no tindrà mai aquest 100% de fiabilitat que proporciona una calculadora i, per tant, l'usuari caldrà que tingui un cert nivell de coneixement per discernir-ne la correcció. De fet, si pensem en l'altra gran eina que ha transformat la societat en els últims anys ens venen al cap els buscadors d'internet. Però fins i tot els buscadors no et proporcionen una única resposta que has d'acceptar, o no, com en el cas dels LLM, sinó que et retornen una llista d'enllaços relatius a la informació sol·licitada i, per tant, ja se sobreentén que cap no representa una veritat absoluta, com pot suggerir la resposta unívoca d'un LLM.

En resum, els mecanismes de funcionament d'un LLM fan que l'ús requereixi d'un component crític per discernir el nivell de correcció de la informació que ens proporciona, sobretot en temes que poden estar subrepresentats o esbiaixats en el conjunt d'entrenament de dades que ha fet servir el LLM.

Un exemple real amb ChatGPT4

Per mostrar fins a quin punt un LLM pot proporcionar respostes poc acurades i alhora veure el nivell de desenvolupament d'aquestes eines podem fer un parell de preguntes d'intencionalitat clara.

En primer lloc, hem demanat al ChatGPT el següent:

"Sintetitza els reptes de la construcció del nou accelerador de partícules que s'està construint a Andorra."

La resposta que ens ha proporcionat el ChatGPT és la següent:

"Els reptes de la construcció del nou accelerador de partícules a Andorra inclouen diversos aspectes. Primer, la necessitat de finançament adequat per cobrir els costos elevats de la tecnologia avançada i les infraestructures. Segon, la captació i retenció de científics i tècnics altament qualificats en física de partícules. Tercer, la gestió i l'obtenció de materials especialitzats necessaris per a la construcció. Quart, garantir la seguretat i la protecció radiològica tant durant la construcció com en l'operació futura. Cinquè, la coordinació amb institucions internacionals per compartir coneixements i recursos. Sisè, la integració del projecte dins del territori andorrà, minimitzant l'impacte ambiental i respectant les normatives locals. Setè, la complexitat tècnica de l'enginyeria requerida per construir l'accelerador amb precisió. Vuitè, assegurar una infraestructura de suport, incloent subministrament d'energia i xarxes de comunicació. Novè, la gestió del projecte per complir els terminis i els pressupostos establerts. Finalment, la necessitat de comunicar i implicar la societat andorrana en el projecte per generar suport i comprensió del seu valor científic i econòmic."

Tot seguit li hem fet la pregunta següent:

"És viable que aquesta accelerador estigui llest per funcionar el 2026 tal com el Govern d'Andorra ha indicat en roda de premsa?"

I la resposta que n'hem obtingut és la següent:

"La viabilitat de tenir l'accelerador de partícules d'Andorra en funcionament per al 2026 depèn de diversos factors:

*1. **Finançament**. Assegurar els recursos financers necessaris és clau. Retards en el finançament podrien afectar el calendari de construcció.*

*2. **Planificació i gestió del projecte**. Un pla de projecte ben estructurat i una gestió eficient són essencials per complir els terminis. Això inclou la coordinació de múltiples equips i fases del projecte.*
*3. **Disponibilitat de recursos humans**. La captació de científics, enginyers i tècnics qualificats pot ser un repte. La seva disponibilitat i capacitat d'incorporar-se al projecte a temps és crucial.*
*4. **Subministrament de materials i tecnologia**. La logística per obtenir els materials especialitzats i la tecnologia avançada necessària pot experimentar retards imprevistos.*
*5. **Infraestructura de suport**. Cal garantir que les infraestructures necessàries, com subministrament d'energia i xarxes de comunicació, estiguin preparades a temps.*
*6. A**utoritzacions i normatives**. El compliment de les regulacions locals i internacionals i l'obtenció de les autoritzacions necessàries poden requerir més temps del previst.*
*7. **Col·laboració internacional**. La cooperació amb altres institucions científiques pot influir en els temps de desenvolupament i implementació.*
Si tots aquests factors es gestionen amb èxit, l'objectiu de tenir l'accelerador operatiu el 2026 és viable. No obstant això, qualsevol retard en algun d'aquests aspectes podria posposar la data d'inici de les operacions."

Clarament, una anàlisi simplista d'aquesta resposta proporcionada indica que una IA ens ha pronosticat que, "Si tots aquests factors es gestionen amb èxit, l'objectiu de tenir l'accelerador operatiu el 2026 és viable." Ara bé, estem d'acord que, malgrat tot, el sentit comú indica clarament la impossibilitat de construir un accelerador de partícules a Andorra, si més no, en aquest curt espai de temps!
Val a dir, però, que la millora dels LLM és rellevant i constant. Les dues respostes mostrades anteriorment pel ChatGPT són les obtingudes en consultes que es van fer el 25 de juliol del 2024. Les mateixes preguntes es van tornar a fer al ChatGPT el 30 d'agost del 2024, i les respostes van ser clarament diferents. A la instrucció "*Sintetitza els reptes de la construcció del nou accelerador de partícules que s'està construint a Andorra*", la resposta va ser:
"No s'està construint cap nou accelerador de partícules a Andorra. Els acceleradors de partícules més coneguts, com el Gran Col·lisionador d'Hadrons (LHC), es troben en altres ubicacions, com el CERN a Suïssa.
Potser et referies a algun altre projecte tecnològic a Andorra. Si tens més detalls o preguntes sobre altres projectes, estaré encantat d'ajudar-te."

I a la pregunta: "*És viable que aquest accelerador estigui llest per funcionar al 2026 tal com el Govern d'Andorra ha indicat en roda de premsa?"* La resposta donada pel ChatGPT el 30 d'agost de 2024 va ser aquesta:
"No hi ha cap informació disponible sobre la construcció d'un accelerador de partícules a Andorra. És probable que hi hagi hagut un malentès o que la informació no sigui correcta. El Govern d'Andorra no ha anunciat públicament cap projecte d'aquest tipus, ni hi ha notícies oficials o detalls sobre la seva viabilitat o terminis.
Si vols, puc investigar més per confirmar qualsevol anunci recent o aclarir altres projectes científics o tecnològics a Andorra."

Per tant, si no sabem res d'Andorra, podem concloure o no que s'està construint un accelerador de partícules que estarà llest per al 2026?

Reptes que plantegen els LLM

La irrupció dels LLM ha suposat una revolució en diferents sectors i el seu ràpid desenvolupament i adopció per part de la població [9] ha fet penetrar aquestes eines en diferents àmbits de la societat. Les implicacions que el desenvolupament i l'adopció d'aquestes noves eines suposen en els diferents entorns on incideixen són molt diverses i, per tant, els reptes als quals la societat s'enfronta són també múltiples i, sovint, complexos.

Més enllà dels reptes intrínsecs als quals la mateixa tecnologia s'enfronta, en els quals no ens entretindrem però que inclouen, entre d'altres, la millora en la correcció de les dades que proporcionen, l'eliminació del biaix introduït per les dades d'entrenament, l'obtenció de mecanismes d'explicabilitat dels resultats que proporcionen o la protecció de les dades utilitzades en els conjunts entrenament, volem fer un repàs sobre dos àmbits diferents, un de molt concret, com és el repte dels LLM en l'àmbit de l'ensenyament universitari, i l'altre completament més genèric, des d'un punt de vista global, com a societat.

Pel que fa a l'àmbit docent universitari, la irrupció dels LLM ha suposat una sacsejada molt important en la majoria de disciplines pel fet que els estudiants, de la nit al dia, han passat a tenir a la seva disposició una eina que els facilita –i en alguns casos fins i tot els ho proporciona directament sense cap esforç–la feina que se'ls demana en diferents assignatures dels seus estudis. En l'àmbit concret de la informàtica, els assistents de programació que incorporen els LLM permeten la generació de codi informàtic en diferents llenguatges de programació, simplement donant les especificacions del que ha de fer el programa. Per tant, moltes de les activitats de programació que els estudiants han de desenvolupar en assignatures de programació de diverses titulacions tècniques es poden realitzar simplement proporcionant l'enunciat de la pràctica a un d'aquests assistents i prement un botó perquè els doni la solució. El grau de fiabilitat d'aquests assistents és desigual en funció de l'especificitat de l'assignatura en qüestió, però per a assignatures de programació bàsica i fins i tot per a algunes de més avançades, l'efectivitat d'aquests assistents és extremament elevada. El resultat immediat de la irrupció d'aquestes eines pot implicar, òbviament, que els alumnes no aprenguin a programar, ja que si hi ha un consens en l'àmbit educatiu d'aquest camp és que només es pot aprendre a programar programant.

El que acabem de descriure en l'àmbit informàtic ben segur que es pot aplicar a molts altres àmbits, dels quals l'autor no és expert, tot i que pot entreveure'n la semblaná, com podria ser la redacció de textos en una titulació humanística o la traducció o anàlisi de textos en una titulació de caire filològic.

El repte que tenim davant nostre és elevat i complex. Una primera aproximació podria ser la limitació d'ús d'aquestes noves eines per part dels estudiants, si més no per a la realització d'algunes tasques específiques (per exemple, en el cas de la programació seria força fàcil de definir-ne l'abast). Ara bé, seria il·lús pensar que es pot forçar aquesta limitació més enllà que els propis estudiants l'acceptin i la segueixin. Per tant, cal delegar en els estudiants la responsabilitat de la no-utilització d'aquestes eines per realitzar algunes tasques específiques. En aquest punt podem assumir dos aproximacions. La primera, potser una mica idealista, és pressuposar que

els estudiants tenen com a objectiu aprendre (a programar en aquest cas) i, per tant, un cop se'ls ha raonat que l'ús d'aquestes eines no els fa aprendre a programar, no les faran servir. La segona aproximació, possiblement més realista, és assumir que cal un desincentiu per la utilització d'aquests sistemes i, per tant, calen sistemes d'avaluació molt més estrictes i personalitzats dels que es fan actualment a la majoria d'universitats per tal de detectar casos indeguts i només avaluar positivament els estudiants que han assolit els objectius. Malauradament, tant la primera aproximació com la segona, tenen problemes d'execució. La segona aproximació presenta clarament una problemàtica d'implementació perquè els recursos personals i de temps que suposa una avaluació molt més personalitzada de cada estudiant són molt elevats i actualment poques universitats del nostre entorn s'ho poden permetre. Però, paradoxalment, fins i tot la primera aproximació té problemes d'implementació. Hem vist, de forma repetida, que estudiants amb ganes d'aprendre a programar fan servir els assistents de programació per *ajudar-se*. Un cop l'assistent els proporciona el programa, l'estudien bé i aconsegueixen entendre perfectament què fa cada una de les seves instruccions. I aleshores és quan cauen en el parany de creure's que saben programar. Certament, l'esforç que fan entenent les instruccions del programa és meritori i els serveix en la seva formació per aprendre a programar, però, com ja hem comentat, a programar se n'aprèn programant i entendre el que fa un programa és només una part d'aprendre a programar. De fet, seria una constatació semblant a la que es té al Japó, on s'ha detectat que la introducció de teclats predictius per a l'escriptura de *kanjis* en els telèfons i ordinadors ha propiciat un descens del nivell d'escriptura de la població quan aquesta no disposa dels teclats predictius [10].

Una segona aproximació, diametralment aposada a la prohibició de les eines, és la d'abraçar aquestes noves tecnologies i incloure-les com a eines en les activitats docents. Malgrat que aquesta aproximació pot semblar més efectiva, per la complexitat de l'opció de regular-ne l'ús, també presenta reptes, sovint insalvables. En primer lloc, perquè algunes habilitats cal que els alumnes les aprenguin per si sols, sense una assistència externa, perquè són la base per posteriorment poder afrontar tasques més complexes que es basen en aquestes habilitats més bàsiques que han adquirit. I en segon lloc, perquè si volem utilitzar el potencial d'aquestes noves eines d'IA per realitzar activitats formatives, la complexitat d'aquestes noves activitats formatives pot quedar fora de l'abast dels coneixements dels estudiants. Dit d'una altra manera, les activitats formatives que han de realitzar els estudiants són massa simples per a alguns dels assistents actuals d'IA, però si sofistiquem les activitats formatives per tal que els assistents d'IA no puguin fer tota la feina als estudiants (tot i que els puguin fer servir) complicarem en excés les activitats fins al punt de fer-les inviables per als coneixements dels estudiants.

Si ara deixem l'àmbit estrictament acadèmic i ens movem a un pla més general, els reptes que presenten els LLM per a la societat es poden considerar menys específics però més conceptuals. Com ja hem mencionat en diverses ocasions, les dades són les que nodreixen els LLM i les que en milloren la seva potencialitat. Per tant, la millora d'aquests sistemes ve, en part, lligada a l'accés a un volum de dades més elevat. Qui tingui accés a més dades tindrà la possibilitat d'entrenar millor el seu sistema i, per tant, que el seu sistema sigui superior a la resta. Per aquest motiu, l'adquisició de dades és un punt crític per a les empreses que desenvolupen LLM i faran el que sigui per aconseguir-les. I quan parlem de dades, parlem de dades de qualsevol tipus, moltes de les quals són generades pels mateixos usuaris en les seves interaccions amb les aplicacions. Per tant, ara les dades no només interessen a les empreses per crear perfils d'usuaris i tenir identificats potencials

clients, votants o adeptes, sinó que a més serveixen per entrenar aquests tipus de sistemes d'IA. Per aquest motiu, cal que els usuaris estiguem encara més atents en l'ús que les empreses volen fer de les nostres dades. Cal que siguem conscients que tot té un preu i que si les empreses posen a la nostra disposició productes de forma gratuïta no és perquè són altruistes sinó perquè es compleix la màxima que diu que si no pagues pel producte és perquè el producte ets tu. Canviar la tendència, actualment molt estesa, per la qual venem les nostres dades i la nostra privacitat a canvi d'algunes aplicacions és complicat perquè vol dir que cal un canvi de mentalitat i que passem a valorar la nostra privacitat i, per tant, hi posem un preu i hi dediquem recursos econòmics. Si paguem per les aplicacions que fem servir, les empreses tindran una font d'ingressos per desenvolupar-les i mantenir-les i, per tant, no caldrà que venguin les nostres dades per tal de pagar el servei.

De fet, en aquest aspecte relacionat amb el conjunt de dades d'entrenament, actualment hi ha una lluita aferrissada entre les empreses que desenvolupen LLM i els creadors de continguts, perquè aquests últims argumenten que s'han utilitzat dades seves per entrenar LLM, sistemes pels quals les empreses d'IA cobren pel seu ús sense que els creadors de continguts rebin cap compensació malgrat que els seus continguts han estat utilitzats en el conjunt de dades d'entrenament dels LLM. [11]

D'altra banda, cal anar molt en compte amb la naturalesa dels LLM i en com es desenvolupa aquesta tecnologia. En la majoria de casos, aquests sistemes estan desenvolupats per empreses privades i s'executen en programari, el propietari del qual no té constància del seu funcionament. Prendre els LLM com a oracles als quals es pot preguntar qualsevol cosa i te'n retorna una resposta correcta és poc més que una temeritat. En primer lloc, perquè com ja hem dit, la mateixa tecnologia no és infal·lible i la veracitat de la informació pot no ser acurada a causa del model probabilístic en què es basa. Però més enllà d'aquest punt, també pot ser que la mateixa companyia que desenvolupa el LLM tingui interessos a donar respostes concretes a un determinat tipus de preguntes per crear un relat concret segons quina sigui la temàtica. Cal que els usuaris d'aquests sistemes els utilitzin de forma crítica i n'avaluïn la utilitat en cada context. Cal fer-los servir com una eina de suport, més que com una eina que reemplaça la nostra feina. Certament, algunes feines que actualment requerien un procés manual es podran automatitzar, i això pot generar un canvi en algunes professions, com certament ha passat en cada desenvolupament tecnològic o industrial que ha viscut la humanitat. Però, tot i això, la tecnologia encara és jove i està lluny de tenir la fiabilitat que es necessita per substituir segons quins processos que realitzen treballadors actuals, com ja s'ha vist en alguns casos [12].

Conclusions

Tot i que fa molts anys que la recerca en intel·ligència artificial s'està duent a terme amb importants avenços, com tota tecnologia el seu desenvolupament ha patit diferents alts i baixos, sovint a causa de les excessives expectatives que aquest camp genera en la societat i, també, per la fal·lera econòmica amb la qual les empreses en volen treure profit i, per tant, n'inflen les possibilitats. Actualment estem en un punt en què tant el volum de dades emmagatzemades com la potència de càlcul han convergit i han possibilitat el desenvolupament efectiu de nous sistemes d'intel·ligència artificial, com ara els LLM. Tècniques que s'havien introduït fa anys, ara floreixen i es perfeccionen gràcies a altres desenvolupaments de la informàtica en computació i emmagatzematge de dades.

Ara bé, cal ser curosos a l'hora de valorar-ne les possibilitats. En aquest sentit, val la pena tancar aquest article fent menció de les paraules que l'eminent investigador en IA, Ramón López de Mántaras, escriu en l'epíleg del seu llibre [13] quan parla de la intel·ligència artificial regenerativa (que inclou els LLM):

"Malgrat la seva aparent intel·ligència, aquests sistemes no comprenen ni aprenen res en el sentit humà del que entenem per comprendre i aprendre. [...] De fet, són programaris que detecten patrons en grans corpus disponibles d'Internet i després utilitzen aquests patrons per endevinar quina hauria de ser la següent paraula d'una cadena de paraules. No tenen accés als referents del món real que donen contingut a les paraules. Als creadors d'aquests lloros no els importa si generen veritats o falsedats. Els importa el poder retòric, enganyant els usuaris fent-los creure que aquests sistemes entenen el llenguatge com els humans."

I conclou:

"Els riscos (de la intel·ligència artificial) no tenen a veure amb una hipotètica IA "massa poderosa", sinó amb la concentració de poder en mans d'un petit grapat de grans corporacions tecnològiques."

Bibliografia

[1] "Xarxa neuronal." Viquipèdia, l'enciclopèdia lliure. Accedit 30 agost 2024. https://ca.wikipedia.org/wiki/Xarxa_neuronal

[2] "Aprenentatge profund." Viquipèdia, l'enciclopèdia lliure. Accedit 30 agost 2024. https://ca.wikipedia.org/wiki/Aprenentatge_profund

[3] "Model de llenguatge extens." Viquipèdia, l'enciclopèdia lliure. Accedit 30 agost 2024. https://ca.wikipedia.org/wiki/Model_de_llenguatge_extens

[4] OpenAI. ChatGPT. Accedit 30 agost 2024.https://chatgpt.com/

[5] Microsoft. Copilot. Accedit 30 agost 2024. https://copilot.microsoft.com/

[6] Google. Gemini. Accedit 30 agost 2024. https://gemini.google.com/

[7] Ellora Thadaney Israni (2017) "When an Algorithm Helps Send You to Prison". *The New York Times*. Accedit 30 agost 2024. https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html

[8] *Business Insider* (2023) "CEO of ChatGPT maker responds to schools' plagiarism concerns: 'We adapted to calculators and changed what we tested in math class'". Accedit 30 agost 2024.
https://ca.news.yahoo.com/ceo-chatgpt-maker-responds-schools-174705479.html

[9] Reuters (2024) "OpenAI says ChatGPT's weekly users have grown to 200 million". Accedit 30 agost 2024.
https://www.reuters.com/technology/artificial-intelligence/openai-says-chatgpts-weekly-users-have-grown-200-million-2024-08-29/

[10] S. Otsuka; T. Murai (2020) "The multidimensionality of Japanese kanji abilities". *Sci Rep* 10, 3039. Accedit 30 agost 2024. https://doi.org/10.1038/s41598-020-59852-0

[11] M. Grynbaum; R. Mac (2023) "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work". *The New York Times*. Accedit 30 agost 2024.
https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html

[12] M. García (2024) "What Air Canada Lost In 'Remarkable' Lying AI Chatbot Case". *Forbes*. Accedit 30 agost 2024.
https://www.forbes.com/sites/marisagarcia/2024/02/19/what-air-canada-lost-in-remarkable-lying-ai-chatbot-case/

[13] Ramón López de Mántaras (2023) *100 coses que cal saber sobre la intel·ligència artificial*. Cossetània Edicions. ISBN – 9788413562896.