

cricc

Centre de Recerca en
Informació, Comunicació
i Cultura



UNIVERSITAT DE
BARCELONA

La intel·ligència artificial generativa en la docència universitària

Jorge Franganillo
Carlos Lopezosa
Marina Salse

Facultat d'Informació i Mitjans Audiovisuals
Universitat de Barcelona
Melcior de Palau, 140
08014 Barcelona

Autors: Jorge Franganillo, Carlos Lopezosa, Marina Salse

Aprovat per la Comissió Acadèmica de Facultat el 27 d'octubre de 2023

Obra distribuïda sota una llicència Creative Commons BY-NC-SA 4.0



Citació recomanada

Franganillo, Jorge; Lopezosa, Carlos; Salse, Marina (2023). *La intel·ligència artificial generativa en la docència universitària*. Barcelona: Universitat de Barcelona.

Sobre els autors

Jorge Franganillo és doctor en Informació i Comunicació, i professor i investigador de la Facultat d'Informació i Mitjans Audiovisuals de la Universitat de Barcelona. Ha desenvolupat activitat docent també a la Hochschule der Medien (Alemanya), a la Universitat de Dubrovnik (Croàcia) i a la Universitat Jan Amos Komenský (República Txeca). És membre del Grup d'innovació docent en Comunicació i Mitjans Audiovisuals (In-COMAV) i de la Càtedra UB DIBA de Comunicació Clara Aplicada a les Administracions Públiques. És també autor dels llibres *Formatos digitales* (2022) i *Gestión de información personal* (2018), i de nombrosos textos acadèmics sobre la gestió de la informació i la comunicació digital. Col·labora com a divulgador a la premsa, on insisteix a remarcar la dimensió humana de la tecnologia.

ORCID: 0000-0003-4128-6546

Contacte: franganillo@ub.edu

Carlos Lopezosa és doctor en periodisme per la Universitat Pompeu Fabra i investigador visitant a la Universitat de Barcelona (beca post-doctoral Margarita Salas). La seva tesi doctoral es va centrar en l'estudi dels factors de posicionament de llocs intensius en continguts i, en especial, de mitjans de comunicació en línia, així com en l'avaluació d'eines d'anàlisi SEO. És especialista en posicionament en cercadors i en sistemes de monetització basats en estratègies de continguts de qualitat. Ha estat professor associat de la Universitat Pompeu Fabra, on ha impartit docència a la Facultat de Comunicació, als graus de Periodisme, Comunicació Audiovisual, i Publicitat i Relacions Públiques.

ORCID: 0000-0001-8619-2194

Contacte: lopezosa@ub.edu

Marina Salse és professora titular d'escola universitària del grau de Gestió d'Informació i Documentació Digital de la Universitat de Barcelona. És llicenciada en Prehistòria, Història Antiga i Arqueologia, i diplomada en Biblioteconomia i Documentació. La seva docència i els seus interessos l'han portada a treballar en àmbits tan diversos com les bases de dades, les metadades, l'anàlisi de contingut o la documentació del patrimoni universitari, àmbits en què compta amb diverses publicacions. És membre del grup d'innovació docent consolidat Adaptabit, dedicat a l'accessibilitat digital de les persones amb discapacitat.

ORCID: 0000-0003-2003-7225

Contacte: salse@ub.edu

Resum

Aquest document té com a objectiu ajudar el professorat universitari a gestionar l'ús la intel·ligència artificial generativa en la seva activitat docent. És un document amb finalitat pràctica que s'inicia amb una presentació del concepte d'intel·ligència artificial (IA) i els seus diversos tipus, després presenta els reptes que la IA suposa per a l'educació superior i finalment analitza la seva aplicació en la docència. En aquest darrer context, s'ofereixen algunes propostes per utilitzar la IA adequadament en l'avaluació de l'alumnat i es proporcionen unes rúbriques que permeten també saber si l'alumnat ha fet un ús correcte d'aquestes eines.

Paraules clau

intel·ligència artificial generativa, educació universitària, avaluació, grans models de llenguatge, *ChatGPT*

Sumari

1. IA generativa i models de llenguatge: una aproximació	4
2. Reptes per a l'educació superior	5
3. IA generativa i aplicacions en la docència	5
3.1. Ús general de la IA per part del professorat	8
3.2. Ús de la IA en els processos d'avaluació de l'alumnat	10
3.2.1. Escenari 1: ús no consentit de la IA per part de l'alumnat	12
3.2.2. Escenari 2: ús consentit de la IA per part de l'alumnat	13
4. Conclusió	14
Referències	16

1. IA generativa i models de llenguatge: una aproximació

La intel·ligència artificial (IA) és un ampli camp de la informàtica que busca crear sistemes capaços d'exercir tasques complexes que fins ara només eren a l'abast de la intel·ligència humana. En l'actualitat, la IA abasta nombrosos subcamps i té aplicacions en moltes àrees, com ara l'aprenentatge, la percepció, el reconeixement de veu, el joc d'escacs, la demostració de teoremes matemàtics, l'escriptura literària o el diagnòstic de malalties (Russell i Norvig, 2004; Boden, 2018; Campesato, 2020; Bhargava i Sharma, 2022).

Una branca d'aquesta tecnologia és la IA generativa, que se centra en la creació de contingut original a partir d'un corpus de dades existent. Aquesta tecnologia fa servir algorismes i xarxes neuronals per «aprendre» d'una gran quantitat d'informació i després generar contingut nou i únic. Un exemple d'aquesta tecnologia és la generació de llenguatge natural (GLN), que és capaç de produir textos en qualsevol idioma. La GLN fa servir grans models de llenguatge com ara GPT, PaLM, LaMDA, LLaMA o Claude, que li permeten «entendre» el llenguatge natural i interactuar amb els humans. La GLN es pot fer servir, doncs, en àmbits diversos, com el periodisme, el màrqueting digital o l'educació, p. ex., per crear resums, informes, notícies, textos creatius o contingut web (Franganillo, 2023).

Els models de llenguatge grans s'han popularitzat a través d'assistents en forma de *xat*, com el *ChatGPT*, el *Bard* o el *xat* de *Bing*, que hi faciliten la interacció. Aquests assistents poden construir text coherent i natural sobre gairebé qualsevol tema, i aquesta capacitat els fa atractius per als qui busquen respostes ràpides i acceptables, encara que no del tot exhaustives o correctes.

La GLN té llavors un gran potencial, però també riscos i reptes relatius a la qualitat i la veracitat del text generat, que pot contenir errors o biaixos perjudicials. Les respostes d'un *xatbot* solen sonar raonables i eloqüents, però poden ser errònies. Tot i que la GLN és una tecnologia versàtil i potent, té les seves debilitats i de vegades produeix contingut inexacte o contradictori. Les respostes proporcionades per un model generatiu no s'han de tractar, doncs, com a informació contrastada i fiable, sinó com a meres construccions sintètiques basades en la probabilitat seqüencial de les paraules, i no en la lògica.

La IA generativa provoca una enganyosa il·lusió de pensament racional, però no raona ni disposa de coneixement fiable sobre el món. No entén, en un sentit humà, res del que escriu (Bender *et al.*, 2021), ni compta amb un model de veritat. Tampoc no està programada per resoldre càlculs o problemes lògics, i falla sovint en tasques que impliquen un raonament d'aquest tipus. Per això, la GLN és vàlida per a situacions que accepten un marge d'error o una certa superficialitat argumental, fins i tot algun disbarat. Però no és apta per a matèries delicades, com ara una consulta mèdica o nutricional, o un consell legal o financer.

La GLN també planteja desafiaments ètics relacionats amb els biaixos que poden afectar els resultats. Els models lingüístics són susceptibles de reproduir els prejudicis i estereotips danyosos que hi ha al corpus usat per entrenar-los. Aquests biaixos semblen avui inevitables; són problemes socials complexos que no es poden resoldre només amb solucions tècniques (Heikkilä, 2023). I impliquen efectes negatius en dos sentits oposats.

En un sentit, poden generar resultats ofensius, inexactes o injustos, com ara afirmacions racistes o sexistes, o poden excloure determinats punts de vista, interessos o grups socials. Però, en l'altre sentit, els biaixos poden fer que la IA sembli poc autèntica, «políticament correcta», fins i tot puritana, cosa que pot limitar la creativitat, la diversitat sociocultural i la llibertat d'expressió si les respostes eviten temes controvertits o censuren paraules considerades polèmiques.

Per totes aquestes raons, el contingut produït per models de llenguatge s'ha de manejar amb cautela i s'ha de sotmetre sempre a una anàlisi crítica.

2. Reptes per a l'educació superior

L'ús de la IA generativa és un tema complex i suposa un gran desafiament, principalment perquè les eines poden retornar respostes de vegades fallides o imprecises, i resulten, en conseqüència, poc fiables. És un problema encara sense resoldre. Estudis recents, com el de Bang *et al.* (2023), confirmen que el *ChatGPT* té una precisió mitjana del 63,41 % en deu categories de raonament específic, la qual cosa implica una alta càrrega d'ambigüitats.

Tot i així, la IA també ofereix beneficis. Malgrat la controvèrsia i les preocupacions que suscita, en poc temps ha atret una notable atenció des de l'àmbit acadèmic (Lopezosa *et al.*, 2023) i des de tota mena d'indústries i sectors empresarials (Ray, 2023).

Així mateix, cal destacar que la Unió Europea ha començat a regular els models d'IA (Llei d'IA, 2023) centrant-se en el que es coneix com el marc de qualitat i el marc d'excel·lència (Comissió Europea, 2020), per bé que és cert que alguns investigadors (Franganillo, 2023; Li, 2023) adverteixen que el paradigma regulatori europeu sobre la IA encara ha d'evolucionar més per mitigar riscos legals i ètics.

Tenint en compte aquestes premisses i que estem davant les fases preliminars del procés regulador europeu, en el següent apartat ens centrem en l'ús ètic de la IA generativa. Per a això, ens recolzem en dos recursos: l'informe de Sabzalieva i Valentini (2023), *ChatGPT e inteligencia artificial en la educación superior: guía de inicio rápido*, publicat per la UNESCO, i la ponència de Nadeu (2023) «Què ensenyem? IA generativa i competències», pronunciada en la jornada *IA i aprenentatge a la universitat*, organitzada per la taula interuniversitària de vicerectorats amb responsabilitats en l'àrea de millora i innovació docent de les universitats catalanes.

3. IA generativa i aplicacions en la docència

Davant els grans reptes que suposa l'ús de la IA generativa com a eina docent en l'educació universitària, el professorat té el compromís d'utilitzar-la de forma responsable. En aquest sentit, és útil la guia de bones pràctiques sobre el *ChatGPT* publicada per la Unesco (Sabzalieva i Valentini, 2023). Aquest document mostra un diagrama de flux dissenyat per Aleksandr Tiulkanov (figura 1), assessor de política de dades i IA, que suposa un bon punt de partida en l'ús responsable del *ChatGPT* i de qualsevol IA generativa.

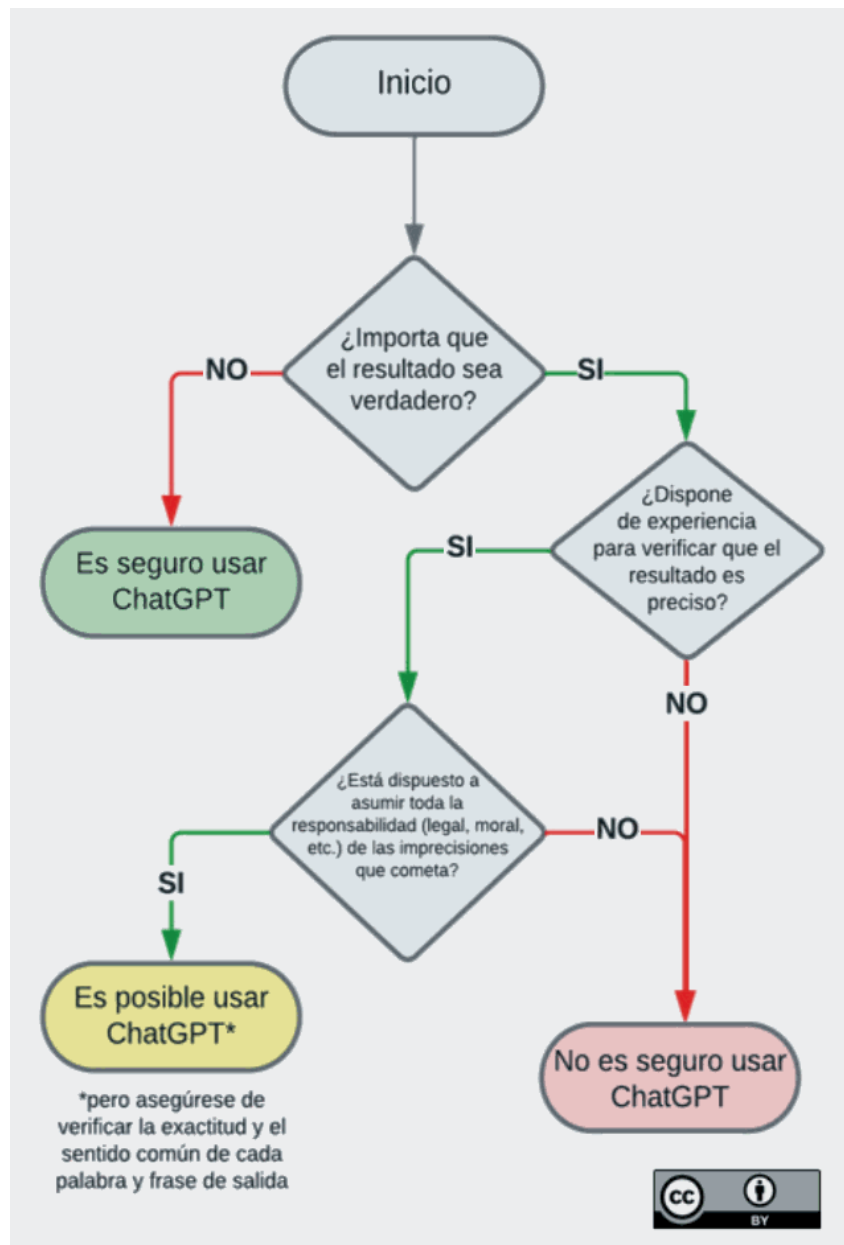


Figura 1. Quan és segur utilitzar el *ChatGPT*?
 Font: Aleksandr Tiulkanov, citat en Sabzalieva i Valentini (2023).

Aquest diagrama mostra la necessitat d'un ús transparent i responsable del *ChatGPT*, que és extrapolable a la formació dels estudiants i dels professors universitaris. Per això, aquesta IA s'ha d'utilitzar sota una doble perspectiva que inclou, d'una banda, l'ètica i, de l'altra, el pensament crític (Lopezosa i Codina, 2023a).

La dimensió ètica es tradueix en el fet que cada context pot tenir una implicació específica que cal atendre, a més que sempre cal ser transparents amb el seu ús i la seva formació. Per tant, els docents hauran d'explicar com es pot utilitzar la tecnologia generativa amb responsabilitat (Rahimi i Abadi, 2023).

Per la seva banda, el pensament crític exigeix reflexionar i verificar la coherència de cada proposta, tant si ve dels docents com dels alumnes, a l'aula o en treballs estudiantils. La guia de la Unesco sobre la IA en l'educació superior (Sabzalieva i Valentini, 2023) és un

referent important per abordar aquest tema des de l'ensenyament i inclou alguns consells. Segons aquest document, la IA pot:

- adaptar el procés d'aprenentatge a les necessitats individuals de l'alumne, fent la seva formació més atractiva i inclusiva;
- automatitzar tasques administratives, oferir suport als professors i fomentar la independència en l'aprenentatge dels estudiants;
- facilitar la creació d'ambients virtuals i personalitzats d'aprenentatge, així com la introducció d'elements de gamificació en l'educació.

El informe també presenta un seguit de recomanacions, gairebé totes elles centrades en l'ètica i en l'ús responsable, per aprofitar el potencial de la IA en la millora de l'educació, com ara:

- desenvolupar polítiques i infraestructures per a l'ús d'IA en l'educació en les quals es promogui un ús responsable i ètic;
- capacitar els docents en l'ús de la IA, perquè puguin aprofitar els seus beneficis a l'aula;
- investigar sobre l'impacte de la IA en l'educació, amb l'objectiu de garantir-ne l'ús segur.

També són rellevants les propostes i recomanacions formulades durant la jornada *IA i aprenentatge a la universitat*. Aquesta jornada va centrar el seu argumentari en el potencial de la IA per transformar l'aprenentatge a la universitat. El discurs es va centrar, igual que l'informe de la UNESCO, en l'ètica i en la responsabilitat. Les idees destacades van ser que la IA:

- té el potencial de revolucionar el procés d'aprenentatge a la universitat en fer-lo més personalitzat, efectiu i eficient;
- té la capacitat d'identificar les dificultats d'aprenentatge dels estudiants i oferir-los un suport adaptat a les seves necessitats individuals;
- pot automatitzar tasques administratives de vegades tedioses, com la correcció d'exàmens o la generació d'informes, permetent que els professors es concentrin més en l'ensenyament;
- ha de fer servir de manera ètica i responsable, evitant qualsevol forma de discriminació cap als estudiants i garantint el seu ús just i imparcial.

Els ponents van proposar integrar la IA en la formació docent, desenvolupar polítiques públiques inclusives i establir marcs ètics per a l'educació amb la IA. Les seves recomanacions van ser:

- incorporar la IA en la formació inicial i contínua dels professors, perquè puguin aprofitar el seu potencial a l'aula;
- desenvolupar polítiques que promoguin la igualtat d'accés i oportunitats en l'educació amb l'ús de la IA;
- establir marcs ètics sòlids que regulin l'aplicació de la IA en l'educació i garanteixin el seu ús responsable i beneficiós per a tothom.

3.1. Ús general de la IA per part del professorat

Els assistents d'IA conversacional, tals com el *ChatGPT*, el *Bard* o el *xat* de *Bing*, poden ajudar a resoldre tasques diverses que requereixen generar un text com a resultat; entre d'altres:

- redacció de textos assagístics o creatius;
- guions de vídeo, de programes de ràdio, de llibres, etc.;
- idees per a presentacions, projectes de recerca, etc.;
- idees per a preguntes d'investigació, redacció d'hipòtesis, etc.;
- comparacions i anàlisi de continguts;
- línies de raonament;
- arguments;
- resums i clarificació de continguts.

Per interactuar amb la IA cal fer servir un *prompt*, que és la petició o instrucció amb què es demana a la IA que faci alguna tasca. Per treure-li el màxim partit a un *prompt*, es recomana seguir aquests consells (Lopezosa i Codina, 2023b):

- detall: com més detallat sigui el *prompt*, millor qualitat tindrà la resposta;
- context: donar-li context al *prompt* és clau perquè la IA entengui millor el que se n'espera;
- objectius: indicar-li a la IA què es vol aconseguir amb el *prompt* ajuda a obtenir respostes més pertinents;
- audiències: assenyalar a qui va dirigit el text pot ajustar la resposta de la IA al context;
- exemples: mostrar exemples relacionats amb la petició ajuda a aclarir el propòsit del *prompt*;
- formats típics d'inici: emprar expressions com «Redacta...», «Compara...», «Crea...», «Resumeix...» o «Actua com...» al començament del *prompt* pot guiar la resposta de manera efectiva.



Figura 2. Full de trucs del ChatGPT. Font: Ana Ivars (2023).

Hi ha, a més, diversos tipus de *prompts* (figura 2). Es poden classificar de la manera següent (Lopezosa i Codina, 2023b):

- reactius: generen interaccions en cadena a partir de cada resposta;
- estructurals: demanen una resposta organitzada en diversos punts, i es demana una ampliació després de cada punt;
- de rol: requereixen que la IA assumeixi un paper específic, com, p. ex., «actua com un expert en...»;
- amb indicació d'audiències: exigeixen que la resposta tingui en compte un públic concret; p. ex., «per a una audiència d'estudiants universitaris»;
- amb objectius: inclouen el propòsit desitjat; p. ex., «per proposar un exercici pràctic sobre recerca booleana...».

Aquests consells pràctics s'han de considerar com a exemples conceptuals de com usar la IA generativa de forma efectiva en diferents interfícies. Tanmateix, cal no oblidar que aquesta tecnologia també es pot integrar en altres aplicacions de productivitat

mitjançant API o altres formes de connexió, la qual cosa pot facilitar l'ús d'aquests *prompts* a gran escala i fer-los més profitosos.

A continuació mostrarem dos possibles escenaris de l'ús de la IA per l'alumnat en els treballs universitaris: un sense el consentiment del professorat i un altre amb el seu consentiment exprés.

3.2. Ús de la IA en els processos d'avaluació de l'alumnat

La irrupció de la IA obliga a revisar els processos d'avaluació de l'alumnat, atès que aquesta tecnologia s'ha nodrit de grans quantitats d'informació i, de forma prou solvent, pot resoldre qüestions de memòria o d'aplicació pràctica, com ara resumir un tema, escriure codi informàtic o consultar una base de dades mitjançant SQL.

En aquest sentit, el professorat hauria de modificar alguns paràmetres per tal de fomentar el raonament crític i reduir l'impacte de la IA en les respostes de l'alumnat. Tot seguit s'assenyalen tres moments clau en què es pot actuar i s'hi indiquen possibles activitats, que s'extreuen principalment de *L'avaluació en temps de la IA* (UNED, 2023)

- **Abans d'iniciar el procés d'avaluació.** En aquest sentit, el centre o el professorat s'han d'ocupar de:
 - Formar l'alumnat en l'ús correcte i ètic d'IA. Tal com s'ha indicat abans, l'alumne ha de ser conscient que les IA cometen errors i que hi ha accions que no són ètiques (com permetre que generi un text i que aquest sigui lliurat sense filtre al professorat). Ha de saber per a què pot fer-lo servir i per a què no.
 - Formar el professorat en l'ús d'IA en la docència.
- **Durant el disseny de les proves.** Referent a això, es podrien explorar algunes idees en els àmbits dels graus de Comunicació Audiovisual (CAV) i de Gestió d'Informació i Documentació Digital (GIDD), ja que tots dos impliquen la IA i permeten treballar el pensament crític.
 - Integrar la IA en els **exercicis**. Es pot proposar als alumnes que li formulin una determinada pregunta a una IA i que avaluïn si la resposta rebuda és correcta.
 - Exemple de CAV. Genera una proposta de guió amb una IA, valora si aquesta proposta és adequada en relació amb la teoria impartida a l'aula i, si no ho és, fes les modificacions oportunes.
 - Exemple de GIDD. Fes amb una IA el resum d'un article i determina si és correcte segons els paràmetres proporcionats pel professor.
 - Treballar els **estudis de cas**. Tot i que la IA pot respondre a determinats casos (especialment, codi informàtic, ja que ha assimilat molts llibres d'informàtica), sovint no podrà respondre adequadament determinades preguntes perquè la informació de què disposa és vaga.
 - Exemple de CAV. Analitza un conjunt d'articles *clickbait*, i indica a quins públics estan dirigits i quin és grau de fiabilitat tenen. Raona la resposta.

- Exemple de GIDD. Quin esquema de metadades podries fer servir per gestionar les fotografies d'un arxiu?
- Realitzar **tasques complexes** o **per fases**, que tinguin molts components i que obliguin a un raonament crític.
- Exemple de CAV. Analitza dues fotografies relacionades amb el turisme a la Cerdanya. Consideres que serien adequades per il·lustrar un web sobre turisme? Quines característiques haurien de reunir? En aquest cas, la IA podrà ser una font de consulta per a l'alumne, que haurà d'aplicar igualment el seu raonament crític en algun dels passos de la tasca.
 - Exemple de GIDD. Elabora diverses recerques sobre IA que proporcionin referències que puguin ser interessants per construir una guia de consulta del tema en una biblioteca pública. Selecciona dues referències, i indica a quin tipus d'usuaris anirien dirigides i per quin motiu.
- Fer **treballs de camp**.
- Exemple de CAV. Grava un vídeo de cinc minuts que reculli l'últim *tardeo* de la Facultat. Edita'l i posa-li subtítols per presentar-lo com a reclam publicitari als nous alumnes.
 - Exemple de GIDD. Visita una determinada biblioteca i compara el que hi has observat amb el que has après a les classes teòriques de la Facultat. Elabora una anàlisi DAFO de la situació d'aquesta biblioteca.
- Emprar **eines de treball** específiques que la IA no pugui utilitzar, tenint en comte que la IA també pot generar gràfics.
- Exemple de CAV. Fes un muntatge audiovisual a partir d'un vídeo proporcionat pel docent, usant el *DaVinci Resolve*.
 - Exemple de GIDD. Dissenya el diagrama entitat-relació d'una botiga virtual d'objectes de regal, usant el *Draw.io* i la notació gràfica apresada a classe.
- **En el moment d'avaluar.**
- Exigir presentacions orals o físiques en què l'alumne hagi de posar en relleu el seu coneixement de la matèria i la seva capacitat de desenvolupament crític.
 - Demanar a l'alumne que indiqui al principi del seu treball si ha fet servir la IA per a aspectes del seu treball, de quina forma i quins mètodes de verificació ha aplicat. Això dependrà del tipus d'exercici, però és aconsellable incloure aquest element en treballs tan decisoris com les memòries de pràctiques o els TFG i TFM.
 - Aplicar les rúbriques i consells que apareixen en [3.2.1](#) i [3.2.2](#) per assegurar l'originalitat i honestedat dels treballs presentats.

3.2.1. Escenari 1: ús no consentit de la IA per part de l'alumnat

El professorat pot detectar si un treball conté text artificial, tot i que no és una tasca fàcil ni infal·lible. No hi ha tàctiques ni eines totalment fiables per identificar contingut artificial, però sí que es poden observar alguns patrons que acostumen a repetir-se.

La rúbrica següent pretén ajudar el professorat a discernir, amb major o menor certesa, si un treball acadèmic podria contenir text creat amb una intel·ligència artificial. Aquesta rúbrica es complementa amb nou eines de detecció de text automàtic, la fiabilitat de les quals —cal insistir-hi— és limitada.

Rúbrica per identificar treballs desenvolupats amb IA	
Comprovacions	Sí/ No
Revisa el text i identifica els patrons més habituals	
El treball lliurat inclou «En resum» i altres col·legues al final del text?	
S'observen repeticions de frases i paraules al llarg de la feina, que el fan sonar superficial o artificioós?	
El treball inclou dades falses o imprecises?	
El treball tendeix a emprar frases curtes amb l'estructura de subjecte + verb + predicat, i manca de paràgrafs llargs?	
Hi ha successions de paràgrafs en forma de tesi + antítesi + síntesi?	
El treball incorpora referències bibliogràfiques falses?	

Després d'aplicar aquesta primera rúbrica, se suggereix copiar el contingut del treball i enganxar-lo en alguna eina d'identificació de contingut artificial. La taula següent enumera una selecció d'aquestes. Per a un bon ús, es proposa provar-ne dues o tres amb el text a avaluar i observar-hi el grau de coincidència.

Eina	Informació general	URL
<i>Content at scale</i>	Genera contingut amb IA i que, al seu torn, compta amb un sistema de detecció de contingut artificial en el qual mostra la proporció de contingut creat per la IA i per l'usuari.	https://contentatscale.ai
<i>Copyleaks</i>	Pot detectar contingut generat per diversos models de llenguatge.	https://copyleaks.com
<i>Corrector App</i>	Ajuda a corregir i millorar textos. Compta també amb una funció d'identificació de contingut artificial. Per tal que funcioni correctament, cal escriure-hi almenys 300 paraules.	https://corrector.app
<i>Crossplag</i>	Detecta contingut artificial. El seu model identifica les parts específiques realitzades per la IA.	https://crossplag.com
<i>GPT Zero</i>	Avalua si un contingut està creat amb IA. Segons recull el seu web, és el detector d'IA amb més volum d'usuaris.	https://gptzero.me
<i>Originality.AI</i>	Extensió de <i>Google Chrome</i> capaç d'identificar si una pàgina web està creada per una IA.	https://originality.ai
<i>Plagium</i>	Detecta contingut plagiat, ja sigui aquest generat o parafrasejat per IA. És una eina basada en GPT.	https://plagium.com/es/ai_detector

<i>Sapling</i>	Només necessita 50 paraules d'un text per identificar si el contingut és d'origen humà o artificial.	https://sapling.ai
<i>Smodin</i>	Parafraseja textos de fins a 1 000 paraules. Compta a més amb un detector de contingut d'IA multilingüe.	https://smodin.io/es/detector-de-contenido-de-ia
<i>Writer</i>	Ajuda a escriure tot tipus de textos i inclou un apartat d'identificació de text creat per mitjans generatius.	https://writer.com

En cas que el professorat observi indicis clars que un treball conté text artificial, es recomana que ho comunicui a l'estudiant i li demani una explicació. També pot sotmetre'l a una prova de validació. Si l'estudiant no aporta proves suficients que el text és original i de la seva autoria, o si es nega a col·laborar, es considerarà que ha comès una irregularitat i se li aplicarà la normativa d'avaluació vigent.

3.2.2. Escenari 2: ús consentit de la IA per part de l'alumnat

Si el professorat autoritza els estudiants a utilitzar eines de GLN als seus treballs, els ha d'ensenyar a fer-ho de forma responsable i ètica (Rahimi i Abadi, 2023). Els alumnes han de fer servir la IA com un ajut, no com un substitut, per crear contingut precís i de qualitat. Per a això, han d'aplicar el pensament crític i la transparència en tot el procés (Lopezosa i Codina, 2023a; Lopezosa *et al.*, 2023), i utilitzar fonts d'informació fiables i verificables.

L'ètica implica considerar les implicacions de cada context i ser transparents sobre l'ús de les eines de GLN. El pensament crític suposa verificar la coherència i la rellevància de cada proposta i reflexionar-hi. La transparència exigeix documentar tot el procés de les peticions a la IA i el seu resultat com a treball dels alumnes, i indicar les fonts d'informació que s'han consultat.

Algunes eines de GLN, com ara el *ChatGPT*, no mostren les fonts d'informació o ho fan amb errors (almenys, fins la data d'aquest informe), cosa que en dificulta la verificació i la transparència. Per això, es recomana que els estudiants facin servir altres eines que sí ofereixen fonts en les seves respostes. La taula següent en mostra algunes:

Eina	Informació general	URL
<i>Bing Chat</i>	Sistema d'IA basat en el model GPT, integrat en el cercador <i>Bing</i> .	https://bing.com
<i>Bard</i>	Sistema d'IA basat en el model PaLM, creat per <i>Google</i> .	https://bard.google.com
<i>Perplexity IA</i>	Cercador conversacional connectat a Internet.	https://perplexity.ai
<i>YouChat</i>	Assistent de recerca d'IA amb el qual pot parlar directament en els resultats de recerca.	https://you.com

L'ús d'eines de GNL pot facilitar el treball acadèmic, però també implica riscos i desafiaments. Per això, estudiants i docents han de ser crítics amb les propostes de treball que es recolzin en aquestes eines. Els estudiants han d'indicar clarament com han fet servir la IA en els seus treballs, quins resultats han obtingut, com els han verificat i validat, i com han aplicat el pensament crític durant el procés (Codina, 2023). I els docents, per la seva

banda, han d'examinar amb cura els lliuraments dels alumnes per detectar possibles errors, biaixos, i fonts poc fiables o falses.

Per ajudar a avaluar l'ús òptim de la IA en els treballs acadèmics, es proposa aquí una rúbrica basada en el diagrama de Codina i Garde (2023), que considera aspectes com ara la verificació, la fonamentació, l'ampliació i l'aprofundiment. A més, s'hi afegeixen aspectes com ara l'ús adequat de les fonts consultades i la transparència (Lopezosa i Codina, 2023a).

Rúbrica per avaluar els treballs desenvolupats amb IA generativa		
Comprovacions	Descripció	Sí/ No
Transparència	El treball declara l'ús d'un assistent d'IA com a part de la metodologia de treball?	
	El treball incorpora un annex amb els <i>prompts</i> i les respostes originals de l'assistent d'IA?	
	El treball explica clarament la raó per la qual es va fer servir l'assistent d'IA?	
Verificació	L'estudiant analitza el resultat del <i>prompt</i> de forma crítica i defensa en un paràgraf si el resultat és correcte o erroni, i per què? En cas que sigui impossible de verificar, l'estudiant explica per quina raó és impossible?	
Fonamentació	L'estudiant identifica les idees més destacades i de valor dels resultats dels <i>prompts</i> utilitzats i justifica la seva decisió?	
	L'estudiant identifica i aplica a fonts de qualitat i d'autoritat externes a l'assistent d'IA per completar el <i>prompt</i> ?	
	Les noves fonts identificades i aplicades estan referenciades en el text i en la bibliografia?	
Ampliació	L'estudiant amplia la proposta conceptual, no només amb les noves fonts identificades, sinó també amb les seves reflexions, opinions o creativitat?	
Aprofundiment	L'ampliació de l'estudiant és reflexiva, adequada i coherent, allunyada de llocs comuns o imprecisions?	

Adaptat de Codina i Garde (2023)

Després d'aplicar aquest procés, que inclou tant l'ús de la IA com l'aplicació de l'ètica, el pensament crític i la transparència, es pot assolir un nou coneixement vàlid i enriquit.

En definitiva, la IA s'ha d'utilitzar en un marc de responsabilitat. El seu ús pot ser un punt de partida, però no ha de substituir l'esforç intel·lectual humà. A més, la IA no pot ser considerada autora d'allò que genera, ja que no té consciència ni responsabilitat, i no se li pot demanar rendició de comptes. Per tant, en usar la IA, s'ha de respectar l'ètica i la transparència, i s'ha d'exercitar el pensament crític.

4. Conclusió

La IA ha arribat per quedar-se. És encara imperfecta, però aprèn ràpid i compta amb constants inversions per a la seva millora contínua. El personal docent no pot ignorar la seva existència, sobretot quan els alumnes tampoc ho faran. Per això, és important que es

formi en IA per comprendre a què s'enfronta i aconseguir que sigui una aliada en els processos d'aprenentatge de l'alumnat.

Alhora, també és important formar els estudiants perquè sàpiguen discernir què és vàlid, i què no, entre les respostes obtingudes, així com educar-los en l'ús ètic de la tecnologia generativa, perquè aprenguin quan és convenient utilitzar-la i quan estan perjudicant amb ella la seva pròpia educació.

Aquest document pretén ser un punt de partida per ajudar el professorat universitari a gestionar millor la relació amb aquest nou convidat que ha arribat a les aules. Esperem que sigui d'utilitat i que contribueixi a fomentar un ús responsable, crític i transparent de la IA generativa en l'àmbit educatiu.

Referències

- Aïmeur, E., Amri, S., Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*, 13(1), 30. <https://doi.org/10.1007/s13278-023-01028-5>
- Alkaissi, H., McFarlane, S.I. (2023). Artificial hallucinations in *ChatGPT*: implications in scientific writing. *Cureus*, 15(2). <https://doi.org/10.7759/cureus.35179>
- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., Do, Q.V., Xu, Y., Fung, P. (2023). A multitask, multilingual, multimodal evaluation of *ChatGPT* on reasoning, hallucination, and interactivity. *arXiv preprint*. <https://arxiv.org/abs/2302.04023>
- Bender, E. (2021). On the dangers of stochastic parrots: can language models be too big? A: *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bhargava, C., Sharma, P.K. (eds.) (2022). *Artificial intelligence: fundamentals and applications*. CRC Press.
- Boden, M.A. (2018). *Artificial intelligence: a very short introduction*. Oxford University Press.
- Campeato, O. (2020). *Artificial intelligence, machine learning and deep learning*. Mercury Learning and Information.
- Codina, L. (2022). *Cómo utilizar ChatGPT en el aula con perspectiva ética y pensamiento crítico: una proposición para docentes y educadores*. <https://lluiscodina.com/chatgpt-educadores>
- Codina, L., Garde, C. (2023). *Uso de ChatGPT en la docencia universitaria: fundamentos y propuestas*. <http://hdl.handle.net/10230/57015>
- Comisión Europea (2020). *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*. Oficina de Publicaciones de la Unión Europea.
- Diakopoulus, N. (2023) Can *ChatGPT* help journalists fact-check faster? *Medium*. <https://generative-ai-newsroom.com/can-chatgpt-help-journalists-fact-check-faster-351e64a2ef32>
- Frackiewicz, M. (2023). *ChatGPT-4* and the future of fake news detection: AI-driven fact-checking and information verification. *TS2*. <https://ts2.space/en/chatgpt-4-and-the-future-of-fake-news-detection-ai-driven-fact-checking-and-information-verification>
- Franganillo, J. (2023). La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos. *Methados: revista de ciencias sociales*, 11(2), m231102a10. <https://doi.org/10.17502/mrcs.v11i2.710>
- Guida, G., Mauri, G. (1986). Evaluation of natural language processing systems: issues and approaches. *Proceedings of the IEEE*, 74 (7): 1026–1035. <https://doi.org/10.1109/PROC.1986.13580>
- Heikkilä, M. (2023). Why it's impossible to build an unbiased AI language model. *MIT Technology Review*. <https://technologyreview.com/2023/08/08/1077403>

- Hoes, E., Altay, S., Bermeo, J. (2023). Using *ChatGPT* to fight misinformation: *ChatGPT* nails 72 % of 12,000 verified claims. <https://doi.org/10.31234/osf.io/qnjkf>
- Li, Z. (2023). The dark side of *ChatGPT*: legal and ethical challenges from stochastic parrots and hallucination. *arXiv preprint*. <https://arxiv.org/abs/2304.14347>
- Lopezosa, C., Codina, L. (2023a). *ChatGPT* y software CAQDAS para el análisis cualitativo de entrevistas: pasos para combinar la inteligencia artificial de OpenAI con *ATLAS.ti*, *NVivo* y *MAXQDA*. Barcelona: Universitat Pompeu Fabra. Departament de Comunicació, 94 p. (DigiDoc Reports).
- Lopezosa, C., Codina, L. (2023b). Inteligencia Artificial y métodos cualitativos de investigación: *ChatGPT* para revisiones de la literatura y análisis de entrevistas semiestructuras. <http://eprints.rclis.org/44296>
- Lopezosa, C., Codina, L., Ferran-Ferrer, N. (2023). *ChatGPT* como apoyo a las systematic scoping reviews: integrando la inteligencia artificial con el framework SALSA.
- Nadeu, C. (2023). Què ensenyem? IA generativa i competències. En: Universitat Politècnica de Catalunya (2023). IA i aprenentatge a la universitat. *YouTube*. <https://youtu.be/oNpPF9U3cVY&t=2258s>
- OpenAI (2022). *Introducing ChatGPT*. <https://openai.com/blog/chatgpt>
- Rahimi, F., Abadi, A.T.B. (2023). *ChatGPT* and publication ethics. *Archives of Medical Research*, 54(3), 272-274.
- Ray, P.P. (2023). *ChatGPT*: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Russell, S., Norvig, P. (2004). *Inteligencia artificial: un enfoque moderno*. Pearson Educación.
- Sabzalieva, E., Valentini, A. (2023). *ChatGPT e inteligencia artificial en la educación superior: guía de inicio rápido*. Unesco. https://unesdoc.unesco.org/ark:/48223/pf0000385146_spa
- UNED (2023). *La evaluación en tiempos de la IA*. <https://canal.uned.es/series/64b4e9cd32e2ca202d026bc5>