Mantenir la integritat acadèmica: els reptes de la intel·ligència artificial generativa

Aleu Miret Moscatel. eUniv
ORCID: 0009-0003-2964-3798

Resum
L'article analitza l'impacte disruptiu de la intel·ligència artificial generativa (IAG), com ChatGPT, en la integritat acadèmica en l'educació superior. A partir d'una revisió exhaustiva de la literatura, es posa de manifest com l'accés generalitzat a aquestes eines pot facilitar conductes deshonestes com el plagi o l'e-cheating, especialment en entorns d'avaluació en línia. S'hi analitzen marcs teòrics com la prevenció del frau, el triangle del frau i els factors individuals i institucionals que afavoreixen la deshonestedat. L'estudi proposa estratègies com la redefinició dels mètodes d'avaluació, l'actualització de normatives, el foment d'una cultura institucional d'integritat i l'ús pedagògic i ètic de la IAG. Finalment, es defensa que l'IAG no ha de ser rebutjada, sinó integrada responsablement en l'ensenyament universitari a través d'un canvi profund en l'avaluació, la metodologia docent i la formació d'estudiants i docents.

Paraules clau
intel·ligència artificial generativa, integritat acadèmica, plagi, avaluació en línia, educació superior, frau acadèmic

Abstract
This article examines the disruptive impact of generative artificial intelligence (GAI), such as ChatGPT, on academic integrity in higher education. Through an extensive literature review, it highlights how widespread access to these tools can facilitate dishonest behaviors such as plagiarism and e-cheating, especially in online assessment contexts. The study analyzes theoretical frameworks like situational crime prevention, the fraud triangle, and individual and institutional factors that foster academic dishonesty. It proposes strategies including the redesign of assessment methods, updates to institutional policies, fostering a culture of academic integrity, and ethical, pedagogical use of GAI. Ultimately, the article argues that GAI should not be rejected but responsibly integrated into higher education through comprehensive changes in evaluation, teaching methodology, and training for both students and educators.

Keywords
generative artificial intelligence, academic integrity, plagiarism, online assessment, higher education, academic dishonesty

IA

1. Introducció

La irrupció de la Intel·ligència Artificial (IA) generativa, especialment amb l'adveniment del ChatGPT el 30 de novembre de 2022, ha provocat un canvi de paradigma en el món educatiu. Un dels aspectes més preocupants d'aquest canvi és el seu impacte en la integritat acadèmica, ja que facilita la producció de treballs plagiats o deshonestos per part dels estudiants. Aquest treball té com a objectiu principal realitzar una revisió exhaustiva de la literatura científica per a ponderar adequadament si la IA generativa ha canviat el pas de la integritat acadèmica i, en cas afirmatiu, quin és el grau d'afectació. Es pretén determinar si es tracta d'un problema realment nou i quines solucions s'han proposat per abordar-lo.

És preceptiu incloure en aquesta introducció una delimitació conceptual dels dos elements que conformen l'enunciat d'aquest treball, això és, la integritat acadèmica i la Intel·ligència Artificial generativa. Pel que fa al primer concepte, agafarem prestada la definició que l'International Center for Academic Integrity (ICAI) estableix i que ha estat replicada aquí i arreu. L'ICAI és una institució fundada l'any 1992 per Don McCabe, professor de la Rutgers University, i que es destaca entre d'altres per haver publicat dues edicions de l'ICAI Reader, un compendi de literatura científica versada en integritat acadèmica que actua com a manual de referència sobre la matèria. Així, l'ICAI defineix integritat acadèmica com el "commitment to six fundamental values: honesty, trust, fairness, respect, responsibility, and courage. By embracing these fundamental values, instructors, students, staff, and administrators create effective scholarly communities where integrity is a touchstone. Without them, the work of teachers, learners, and researchers loses value and credibility. More than merely abstract principles, the fundamental values serve to inform and improve ethical decision-making capacities and behavior. They enable academic communities to translate ideals into action" (ICAI, 2021). Per tant, la integritat acadèmica és una expectativa que ha de ser compartida per tots els seus agents de cara a fomentar un clima sa de desenvolupament educatiu i, en conseqüència, ha de ser invocada, informada i discutida de forma regular.

En afegit, és necessari desenrotllar amb un cert detall la tipologia d'actes perpetrats pels estudiants i que erosionen aquest clima anhelat d'integritat acadèmica[1]. La forma més comuna de conducta deshonesta és el plagi, entès com l'entrega d'una tasca o treball que no és propi i no se'n referencia l'autoria. El coneixement o no d'aquesta mancança per part de l'alumne no eximeix de la falta, ni tampoc el seu abast - si és una còpia literal o una paràfrasi -. Existeixen, no obstant, molts altres actes que atempten contra la integritat acadèmica, com per exemple el reciclatge de treballs propis ja entregats i avaluats, la invenció de dades de recerca qualitativa o quantitativa - o qualsevol invenció d'informació, evidència o referència, de fet -, la connivència il·legítima amb altres estudiants per presentar treballs formulats per ser desenvolupats de forma individual, l'engany en un examen (anotacions ocultes, ús de tecnologies fraudulentes, còpia d'altres estudiants, etc.) o la suplantació, sia en un examen o en un entregable i sia mitjançant una contraprestació o no. Com és lògic, podríem trobar altres tipus de desviacions de conductes íntegres, però per aquesta investigació en concret ens interessa centrar-nos sobretot en el plagi, que és en aparença el més afectat per l'estat de desenvolupament i facilitat d'accés actual de la IA generativa.

Pel que fa a aquest darrer concepte, ens remetrem a la definició que ens proporciona el Cambridge Dictionary, per al qual una IA generativa és "a particular artificial intelligence (= a computer system that has some of the qualities that a human brain has, such as the ability to interpret language, recognize images, and learn from data supplied to it) that is able to produce text, images, etc.". Caldria afegir que aquesta capacitat productiva que es posa de relleu en aquesta definició s'assoleix a través de l'enginyeria *prompt* i a partir d'un entrenament, això és, d'unes dades d'entrada que són assimilades per a reproduir-ne de similars a partir de l'estructura i patrons originals.

[1] Per aquest apartat hem pres com a referència el document *Academic integrity police* (The University of Sydney, 2022), però trobem classificacions idèntiques o molt similars en moltes altres universitats. Alguns exemples són els documents normatius *Statutes and ordinances* (Cambridge, 2022) o Academic Misconduct: Cheating, Plagiarism, and Other Forms (UC Berkeley, 2016).

2. Fonamentació teòrica o conceptual

Si bé la IA troba el seu moment germinal als anys 50 del segle passat, no és sinó en els darrers quinze anys que l'aprenentatge profund de les computadores ha accelerat una evolució des dels models de dades discriminatius a models de dades generatius, permetent saltar de la classificació a la generació d'imatges, textos i audiovisuals (Cao et al., 2023). En aquest context, destaca la fundació l'any 2015 d'OpenAI com a organització sense ànim de lucre orientada a estendre el desenvolupament de la IA generativa en benefici de tota la humanitat. Aquesta organització, que comptava amb figures destacades del món tecnològic dins la pròpia junta directiva - tals com Elon Musk o Sam Altman - i amb el finançament de Microsoft i Amazon, entre molts d'altres, ha marcat el pas del progressiu desenvolupament de models de llenguatge extens (LLM, per les sigles en anglès de *large language model*). Així, si el 2018 llançaven el GPT-1 (*Generative Pre-trained Transformer-1*), l'any següent el seguia el GPT-2, el 2020 el GPT-3 i l'any 2022 el GPT-3.5. Si bé en aquest punt ja hi havia unes poques veus que alertaven sobre les implicacions acadèmiques que la IA generativa podia arribar a ocasionar (Abd-Elaal, 2019)academic cheating occurs among both undergraduate and postgraduate students. Consequently, written essays and articles undergo certain detection measures and most teaching and research institutions use a range of software to tackle plagiarism. However, the state-of -art artificial intelligence (AI, no és fins la incorporació d'un *chatbot* a GPT3.5., un software que simula mantenir una conversa amb l'usuari, que la popularització de la IA generativa assolí una magnitud d'abast mundial. El llançament del ChatGPT-4 al març de 2023, de ChatGPT-4o el proppassat 13 de maig de 2024 i tots els desenvolupaments ulteriors que se n'han derivat presenten un escenari d'incertesa i preocupació en múltiples esferes de la societat, mentre que d'altres s'ho miren amb optimisme i entusiasme.

L'adveniment de la Intel·ligència Artificial (IA) generativa d'accés obert ha generat en un any i mig un ingent volum de literatura acadèmica sobre els seus efectes en l'educació superior, que s'ha d'afegir a un no menys ingent volum d'investigacions científiques sobre IA i educació previs a novembre de 2022, que obeïen a un marc conceptual diferent però encara útil – en són exemples els articles de Incio Flores et al. (2021) o García-Peña et al. (2020),

centrats a destacar les oportunitats potencials de la IA en educació més que en ressaltar-ne els riscos -. En concret, i per no perdre'ns entre aquest oceà de referències, ens interessen per a aquesta investigació tots aquells que tractin la interacció entre IA i integritat acadèmica.

Com bé sabem, el fenomen de la integritat acadèmica es discuteix des dels mateixos orígens educatius i és, per tant, molt anterior a l'aparició de la IA o d'Internet. En aquest sentit, constatem que sobretot en la literatura grisa que versa sobre IA i educació l'aproximació que es fa de la integritat acadèmica és molt genèrica i simplificada, i una revisió dels debats sobre integritat acadèmica previs a la IA ens poden aportar una riquesa conceptual que ara mateix no és al centre del tauler. Com segur que hom pot endevinar, el volum de literatura científica sobre integritat acadèmica és ingent i sobrepassa per molt l'espai que ha d'ocupar en aquesta investigació. Només per fer-nos-en una idea aproximada, la segona edició de l'ICAI Reader (2022) abans esmentat, que procura aglutinar els articles de recerca basilars en la matèria, conté 85 documents en total només per al període de 1992 a 2020. Per tant, si bé considerem imprescindible una mirada a les investigacions tradicionals sobre la disciplina, aquesta mirada resultarà per força parcial i incompleta.

És important reconèixer que l'accés a Internet i les TIC a finals dels anys 90 del segle passat ja havia advertit sobre una dinàmica creixent segons la qual les infraccions comeses se sostenien com més va més en una o diverses eines informàtiques i la rapidesa en la preparació i execució del frau era cada cop major (Mut, 2010). Aquesta dinàmica creiem que s'ha accentuat encara més en el període posterior a l'accés a les eines d'IA generativa. Per tant, d'entrada cal avançar algunes característiques sobre integritat acadèmica ja posades de manifest per la literatura prèvia a novembre de 2022. Així, en primer lloc, s'ha manifestat que incórrer en frau acadèmic és fins a cert punt habitual (Curtis i Vardenaga, 2016), amb la majoria dels estudiants cometent-ne d'alguna manera almenys alguna vegada. En segon lloc, igual que amb altres formes de desviació, la major part del frau acadèmic es trobarà a l'extrem menys greu del ventall, com ara no fer una paràfrasi correctament o no citar correctament, en lloc de l'extrem més preocupant, com ara presentar una tasca totalment escrita per un altre o fer que algú es faci passar per tu en un examen (Bretag et al., 2019). En tercer lloc, el frau acadèmic no es distribueix de manera

aleatòria entre els tipus d'avaluació, amb estudiants que fan més trampes en alguns tipus d'avaluacions (com ara les tasques a casa, els exàmens sense límit de temps i els qüestionaris en línia sense supervisió) que en altres (com ara els exàmens vigilats en persona i sense llibres) (Bretag et al., 2019). En darrer lloc, a mesura que augmenta la freqüència d'oportunitats adequades per involucrar-se en frau acadèmic, també ho fa la probabilitat que els estudiants aprofitin aquestes oportunitats (Hodgkinson et al., 2016).

D'aquesta manera, aquesta literatura ja dibuixa un escenari sobre el qual poder dirigir algunes accions que facin front als reptes de la IA, tals com dur a terme avaluacions supervisades de l'estil d'un examen presencial, que ja han demostrat minimitzar els problemes de deshonestedat acadèmica (Bretag et al., 2019), malgrat no totes les competències que s'espera que un alumne desenvolupi poden avaluar-se a partir d'un examen i, en qualsevol cas, aquesta opció no és de fàcil aplicació per a les universitats que ofereixen una formació a distància. També hem vist com es posa l'accent a tractar de reduir al màxim les oportunitats que tenen els estudiants per a incórrer en una acció deshonesta, és a dir, tractar de fer esforços en la prevenció.

Seguint aquesta línia, hi ha força literatura (Awdry & Ives, 2021; Hinman, 2002; Hodgkinson et al., 2016) que ha relacionat la integritat acadèmica i la teoria de l'elecció racional - prou popular en el camp de les ciències econòmiques i socials -, segons la qual els individus tendeixen a maximitzar la utilitat i el benefici i minimitzar els costos i els riscos. Posant-ho en el context actual, si les recompenses associades amb l'ús de la IA per completar un element d'avaluació específic superen els riscos i esforços de fer-ho sense assistència tecnològica, alguns estudiants podrien optar per utilitzar LLMs quan alternativament podrien haver utilitzat una estratègia de frau tradicional, o potser no haurien comès cap frau en absolut.

2.1. *Situational crime prevention*

Anant una mica més enllà, Birks i Clare (2023) proposen un seguit d'estratègies basades en el concepte de *situational crime prevention* (STP) adaptades, aquí sí, a un entorn dominat per l'ús extensiu de la IA generativa. L'STP és una teoria criminològica que es focalitza en el context que proporciona les oportunitats de perpetrar un crim i no pas en l'individu que el comet. En concret, l'STP opera a partir de cinc mecanismes que apel·len cadascun d'ells a problemàtiques específiques del crim: augmentar el risc i l'esforç de l'acte criminal, reduint-ne la recompensa i la incitació, i eliminant les excuses. Aquests autors consideren que és possible augmentar el risc percebut de l'exposició a curt termini de l'enviament de treballs que els estudiants no han fet ells mateixos de diverses maneres. En primer lloc, seria possible automatitzar la comparació entre elements d'avaluació a nivell d'estudiant individual dins d'un mateix curs per veure si hi ha diferències de rendiment inusualment grans entre les avaluacions supervisades (baix risc d'ús de la IA) i les no supervisades (alt risc d'ús de la IA). Diferències grans podrien desencadenar el requisit de defenses de tipus viva - que és com s'anomenen les defenses orals parcials o completes en el món anglosaxó - dels treballs no supervisats. Relacionat amb això, els coordinadors de curs podrien iniciar defenses orals aleatòries de tasques no supervisades de bona qualitat, proporcionant un desincentiu per fer trampes i obtenir bons resultats. Una altra estratègia a curt termini per augmentar el risc percebut de l'enviament de treballs fets per la IA seria la generació d'oportunitats de "denunciants", de manera que els estudiants poguessin denunciar els seus companys que creguin que estan cometent frau acadèmic. En una direcció similar, també proposen explorar la utilitat dels entorns d'aprenentatge d'avaluació en línia que proporcionin mètriques detallades d'ús de l'usuari en la construcció de l'avaluació, com ara copiar i enganxar, eliminació de paraules, patrons d'edició, temps passat en l'entorn d'avaluació, etc.

Una altra mesura interessant desenvolupada per Birks i Clare (2023) és l'augment del risc percebut de la detecció futura, que els autors comparen amb el dopatge esportiu, on s'ha reconegut la possibilitat que els mètodes de detecció millorats en temps futur puguin identificar una mala pràctica que amb la tecnologia present ha passat desapercebuda, motiu pel qual emmagatzemen les mostres dels esportistes per a anàlisis futures. Perquè això funcioni en l'àmbit

de l'educació superior, les universitats haurien de desar totes les avaluacions basades en text enviades pels estudiants al llarg de la seva carrera escolar o universitària. A mesura que es desenvolupin nous detectors de IA, aquestes bases de dades poden ser recorregudes per algorismes automatitzats que apliquin els nous detectors al seu corpus, marcant les avaluacions que anteriorment havien passat desapercebudes però que ara són capturades per nous enfocaments. Una bona comunicació pública d'aquesta pràctica podria actuar com a mesura dissuasiva i reequilibrar una més que probable estimació baixa del risc de detecció actual.

Finalment, pel que fa a l'eliminació d'excuses, Birks i Clare (2023) consideren que és peremptòria una política universitària precisa i inequívoca sobre què és el frau acadèmic i com serà tractat, però no només això: el disseny de les avaluacions fetes pels coordinadors d'estudis, una formació adequada i altres mesures de reforç positiu són igualment imprescindibles. Combinades amb el conjunt d'estratègies relacionades amb el risc/esforç esmentades anteriorment, també és important que les universitats facin complir els procediments de frau sempre que sigui possible i que publicitin (d'una manera que mantingui la privacitat de l'estudiant) que aquest procés d'aplicació de la normativa es produeix. Els estudiants també poden rebre recordatoris específics que alertin la seva consciència, a través de tècniques com ara declaracions d'autenticitat abans de presentar elements d'avaluació no supervisats per a la qualificació.

2.2 *E-cheating* i *e-dishonesty*: la integritat acadèmica en els entorns online

Els mètodes efectius per promoure la integritat acadèmica han estat considerats durant molt de temps dins de l'educació superior. No obstant això, existeix una creença generalitzada que les desviacions de la integritat van començar a incrementar-se amb la introducció de la tecnologia a l'aula i la popularitat de les classes en línia, creant noves oportunitats per al *e-cheating* (Holden et al., 2021). La pandèmia de la COVID-19 ha ocasionat canvis generalitzats en l'educació superior, resultant en l'adopció per part de moltes institucions

de formats d'aprenentatge en línia, sigui de forma completa o parcial, i que s'han mantingut en major o menor mesura un cop els estadis més acusats de la pandèmia ja s'han deixat enrere - de la mateixa manera com el teletreball s'ha implementat en molts sectors laborals pràcticament de forma irreversible -.

En definitiva, ja prèviament a les primeres albors de la IA generativa els professors i l'alta direcció acadèmica van haver d'enfrontar-se al repte de desenvolupar mètodes per avaluar adequadament l'aprenentatge dels estudiants en un entorn en línia mentre es mantenia l'honestedat acadèmica.

Algunes de les trampes pròpies de l'entorn de cursos en línia incloïen, però no es limitaven, a descarregar treballs d'internet i reclamar-los com a propis, utilitzar materials sense permís durant un examen en línia, comunicar-se amb altres estudiants a través de la xarxa per obtenir respostes o fer que una altra persona completi un examen o tasca en línia en lloc de l'estudiant que hi posa nom. En general, hi ha prou evidència per afirmar que tant els professors com els estudiants perceben que les proves en línia ofereixen més oportunitats de fer trampes que en entorns tradicionals amb vigilància en viu, essent les principals preocupacions la col·laboració il·lícita entre estudiants i l'ús de recursos prohibits durant l'examen (Holden et al., 2021).

En una línia similar, el vocable *e-dishonesty* s'ha utilitzat per referir-se als comportaments que s'allunyen de la integritat acadèmica en l'entorn en línia, plantejant noves consideracions que potser no havien estat prèviament considerades pels agents implicats. Per exemple, les preocupacions en relació als exàmens en línia típicament inclouen la manipulació de l'ordinador portàtil o del sistema de gestió de proves, la suplantació d'identitat, la filtració d'ítems d'examen i l'ús de recursos no autoritzats com buscar a internet, comunicar-se amb altres mitjançant un sistema de missatgeria electrònica, comprar respostes a altres, accedir a l'emmagatzematge local o extern de l'ordinador, o accedir directament a un llibre o apunts (Wahid et al., 2015).

Sense que pugui sorprendre a ningú, els factors identificats per la literatura científica que motiven una actitud deshonesta en un entorn *online* són pràcticament els mateixos que els acadèmics van identificar en entorns d'ensenyament presencial i que han estat exposats més amunt. Així, la recerca basada en

el que es coneix com el "triangle del frau" proposa que, perquè es produeixin trampes, han de presentar-se tres condicions: 1) oportunitat, 2) incentiu, pressió o necessitat, i 3) racionalització o actitud (Becker et al., 2006). Totes tres condicions són factors predictius positius del comportament de fer trampes dels estudiants. L'oportunitat sorgeix quan els estudiants perceben que tenen la capacitat de fer trampes sense ser atrapats; aquesta percepció pot ocórrer, per exemple, si es pensa que els instructors i els administradors estan ignorant les trampes evidents o si els estudiants veuen altres fer trampes o reben confidències d'altres estudiants. La segona condició, incentiu, pressió o necessitat, pot provenir de diverses fonts com el propi estudiant, els pares, els companys, els ocupadors i les universitats. La pressió que senten els estudiants per obtenir bones notes i el desig de ser vistos com a exitosos poden crear l'incentiu per fer trampes. Finalment, la racionalització del comportament de fer trampes pot ocórrer quan els estudiants veuen les trampes com a coherents amb la seva ètica personal i creuen que el seu comportament està dins dels límits de la conducta acceptable.

Relacionat amb les factors individuals, s'han delimitat també els factors institucionals, destacant-ne el concepte de la "cultura de les trampes", per la qual la racionalització d'actuar en contra de la integritat acadèmica pot ocórrer si els estudiants creuen que altres estudiants estan fent trampes, detecten una competència injusta o perceben una acceptació o indiferència davant d'aquests comportaments per part dels avaluadors (Tolman, 2017). Els estudiants configuren directament la cultura de fer trampes, i així, subgrups d'estudiants dins d'una població universitària poden tenir les seves pròpies cultures de fer trampes. És plausible, per tant, que els estudiants en línia tinguin la seva pròpia cultura de fer trampes que difereix de la resta de la població estudiantil.

2.3 Mètodes per a reduir la deshonestedat acadèmica en les avaluacions en línia

De la mateixa manera que els motius pels quals els estudiants fan trampes són variats, també ho són els mètodes per reduir la deshonestedat acadèmica. Organitzarem aquest apartat en relació amb els factors propis de quatre dimensi-

ons, que són els habitualment enumerats en la literatura que ha tractat aquesta temàtica (Akbulut et al., 2008; Holden et al., 2021; Tolman, 2017): l'estudiant individual, la institució, el mitjà de lliurament i les avaluacions en si. Al llarg d'aquest epígraf ens centrarem principalment en les avaluacions sumatòries - aquelles que examinen els resultats finals del procés d'aprenentatge, en contraposició a les avaluacions formatives que examinen el propi procés d'aprenentatge -, les quals poden tenir diversos formats, des de preguntes de selecció múltiple fins a proves amb llibre obert. Tot i que els mètodes per prevenir les trampes es discuteixen per separat dels motius pels quals els estudiants fan trampes, subratllem que els mètodes s'han de considerar conjuntament amb els motius i les motivacions que poden portar els estudiants a participar en deshonestedat acadèmica en primer lloc.

2.3.1 Mètodes a nivell individual i institucional

La literatura consultada - anteriorment referenciada - que ha analitzat el fenomen específic dels entorns en línia aglutina tant els mètodes a nivell individual com els de nivell institucional per reduir la deshonestedat acadèmica per influència bidireccional i, per tant, nosaltres seguirem aquest mateix patró. La deshonestedat acadèmica representa un desafiament complex que requereix una resposta integral, que inclogui tant mesures individuals com institucionals. A nivell institucional, un pas fonamental és l'establiment de polítiques clares i detallades sobre què constitueix deshonestedat acadèmica i les conseqüències associades. Aquestes polítiques han de ser comunicades efectivament a tots els membres de la comunitat educativa, incloent estudiants, professors i personal administratiu. La clarificació d'aquestes polítiques ajuda a eliminar la confusió i redueix la probabilitat que els estudiants cometin actes deshonestos sense adonar-se'n.

Els estudis referenciats han demostrat que la percepció de la severitat de les sancions està inversament relacionada amb la freqüència de la deshonestedat acadèmica. Això significa que quan els estudiants creuen que les sancions per fer trampes són severes, són menys propensos a participar en aquestes activitats. Per tant, és essencial que les institucions no només estableixin sancions

clares, sinó que també assegurin que aquestes sancions s'apliquin de manera consistent. La consistència en l'aplicació de les sancions també ajuda a construir una cultura de confiança i integritat dins de la comunitat educativa.

A més, les institucions poden adoptar codis d'honor que defineixin clarament els comportaments ètics i no ètics. Els codis d'honor són efectius perquè augmenten la consciència dels estudiants sobre les normes d'integritat acadèmica i redueixen la seva capacitat de racionalitzar els comportaments deshonestos. Els estudis han trobat que les universitats amb codis d'honor tenen nivells més baixos de deshonestedat acadèmica que aquelles sense aquests codis. Els codis d'honor també poden augmentar la probabilitat que els professors i els estudiants denunciïn les infraccions, contribuint així a mantenir un entorn acadèmic més honest.

2.3.2 Mètodes en relació amb el mitjà de lliurament

Els mètodes relacionats amb el mitjà de lliurament se centren en com s'administren i es vigilen les avaluacions en línia. Una opinió comuna és que els exàmens sumatoris supervisats en persona en centres educatius són els més segurs per mantenir la integritat acadèmica. Tot i això, aquesta opció pot no ser pràctica per als estudiants que viuen lluny del campus o en regions on la supervisió presencial no és factible, com va quedar palès en el context de la pandèmia de COVID-19.

Per fer front a aquests desafiaments, s'han desenvolupat diversos mètodes de detecció de trampes en exàmens en línia que no requereixen supervisió presencial. Aquests mètodes inclouen la supervisió en línia en directe, la gravació de vídeo web i el reconeixement biomètric mitjançant intel·ligència artificial. La supervisió en línia en directe implica que un supervisor vigila els estudiants a través de la seva càmera web durant tot l'examen. Aquesta forma de supervisió és la més similar a la supervisió presencial i pot dissuadir eficaçment les trampes, si bé hi ha una extensa literatura sobre com interacciona aquest mètode amb el dret a la intimitat i la protecció de dades. No entrarem a

valorar-la aquí, si bé volíem deixar constància de les tibantors legals i normatives que se'n poden derivar.

D'una forma similar, la gravació de vídeo web consisteix a registrar els estudiants mentre fan l'examen, amb la possibilitat que els enregistraments siguin revisats més tard per detectar qualsevol activitat sospitosa. Tot i que aquest mètode pot ser útil, pot no ser pràctic revisar tots els enregistraments de manera detallada.

Finalment, el reconeixement biomètric mitjançant intel·ligència artificial utilitza algorismes per detectar comportaments sospitosos durant l'examen - tals com desviar la mirada de la pantalla i el teclat, aixecar-se de la cadira, parlar amb tercers o fins i tot descobrir irregularitats en el patró de tecleig - i marcar aquests moments per a la seva revisió posterior. Aquesta tècnica pot reduir la càrrega de treball dels supervisors, però també pot deixar passar alguns comportaments deshonestos que no segueixen patrons previsibles o que escapen al control biomètric.

Tal i com s'ha advertit, aquests mètodes de supervisió en línia plantegen diverses preocupacions ètiques, incloent la invasió de la privacitat dels estudiants i la seguretat de les dades. A més, hi ha informes de discriminació per part del programari de supervisió en línia, especialment en funció del color de la pell dels estudiants (Swauger, 2020). Per tant, és important que les institucions considerin aquesta àmplia casuística abans d'implementar els mètodes esmentats i garanteixin que es respectin els drets i la privacitat dels estudiants.

2.3.3 Mètodes basats en l'avaluació

Els mètodes basats en l'avaluació se centren en el disseny i la presentació dels exàmens per reduir les oportunitats de fer trampes. Una estratègia és modificar el format de l'examen per fer més difícil la col·laboració no autoritzada i l'ús de materials externs. Per exemple, els exàmens *online* poden dissenyar-se per presentar una pregunta a la vegada, mantenir-la activa durant un temps predeterminat i no permetre que els estudiants tornin a les preguntes ante-

riors un cop les hagin respost. Aquesta tècnica pot limitar la capacitat dels estudiants de compartir preguntes i respostes amb altres durant l'examen o de perpetrar altres estratègies irregulars.

La randomització de les preguntes i de les opcions de resposta és una altra tècnica efectiva. Això significa que cada estudiant rep una versió lleugerament diferent de l'examen, la qual cosa fa més difícil que els estudiants col·laborin entre ells. A més, l'ús de preguntes que requereixin una comprensió conceptual profunda en lloc de la mera memorització de contingut pot reduir la temptació d'utilitzar materials no autoritzats, ja que aquestes preguntes són més difícils de buscar en línia o en altres recursos.

Una altra estratègia és limitar la finestra de temps durant la qual els exàmens estan disponibles. Els exàmens que es poden fer només durant un període de temps curt redueixen les oportunitats de col·laboració entre els estudiants i l'ús de materials externs. També es pot utilitzar programari que bloquegi les funcions de copiar i enganxar durant l'examen, així com navegadors de bloqueig que impedeixin als estudiants accedir a altres llocs web o aplicacions mentre fan l'examen.

A més dels canvis en el format i la presentació dels exàmens, el software que verifica l'originalitat del text (com Turnitin) pot ser útil per detectar el plagi. Aquest tipus de programari compara els treballs presentats amb una base de dades de treballs existents per identificar similituds. Tot i que l'ús d'aquest programari pot ajudar a detectar el plagi, també planteja preocupacions ètiques, com la infracció dels drets d'autor del treball dels estudiants.

Els procediments de control d'exàmens en línia (OECP), o alternatives no supervisades, també poden ser efectius per promoure l'honestedat acadèmica. Aquests procediments inclouen oferir exàmens en un moment determinat, oferir l'examen durant un període breu de temps, randomitzar la seqüència de preguntes, presentar només una pregunta a la vegada, dissenyar l'examen per ocupar un període limitat de temps, permetre l'accés a l'examen només una vegada, requerir l'ús d'un navegador de bloqueig i canviar almenys un terç de les preguntes de l'examen cada trimestre. Aquests mètodes no eliminaran completament les trampes, però poden ajudar a reduir-les significativament.

En resum, en els anteriors paràgrafs s'ha pogut veure com la literatura científica que va ocupar-se de la integritat acadèmica prèvia a l'aparició de la IA generativa, fos en entorns educatius tradicionals o en línia, coincidia majoritàriament en la detecció dels factors que induïen a cometre actes fraudulents i deshonestos. De la mateixa manera, combinaven de forma molt similar els mètodes - a nivell individual, institucional, relacionats amb el mitjà de lliurament i basats en l'avaluació - que havien de contribuir a reduir la deshonestedat acadèmica en les avaluacions de manera efectiva. Aquesta combinació havia de considerar les raons i motivacions dels estudiants per fer trampes i ajustar-se a les necessitats específiques de cada context educatiu. Malgrat puguin extreure's lliçons útils que poden aplicar-se en un entorn educatiu dominat per la IA generativa - i que desenvoluparem en posteriors apartats -, aquesta ha estat prou disruptiva com per convertir algunes de les mesures o eines proposades - com Turnitin - en obsoletes o insuficients.

2.4 Pot la IA generativa ser detectada?

Creuant la frontera de novembre de 2022, observem literatura científica orientada a qüestions purament tècniques sobre la detecció de documents generats per IA. Així, Sadasivan et al. (2024), Liang et al. (2023) o Weber-Wulff et al. (2023) s'ocupen d'alertar-nos sobre les dificultats tecnològiques per a garantir una detecció eficaç i no esbiaixada dels textos generats per IA, fent-nos descartar pel moment una solució tècnica al problema presentat.

Així, deixant de banda les preocupacions relacionades amb la privacitat, les qüestions ètiques associades amb la càrrega del treball de l'estudiant a llocs web de tercers i les qüestions de propietat una vegada el contingut ha estat processat, una gamma de companyies de programari (incloent-hi aquelles que van desenvolupar els models originals) han buscat desenvolupar detectors de text generats per IA. Aquestes aproximacions utilitzen aprenentatge automàtic per analitzar el text i buscar els signes reveladors de text generat per IA, procurant aprofitar dues mètriques de text lliure: *perplexity*, una mesura de la complexitat del text; i *burstiness*, una mesura de la uniformitat de la longitud de les frases. No obstant això, aquests detectors segueixen sent imperfectes en

termes de falsos positius i falsos negatius.

A tall d'exemple definitori, si es fa un cop d'ull al web d'OpenAI podem resseguir com al gener de 2023 es va iniciar un projecte per a desenvolupar un *classifier* (un algorisme que categoritza les dades per etiquetes) que pogués detectar textos generats per la seva pròpia IA generativa i com aquest projecte va ser abandonat al juliol del mateix any per la seva poca precisió[2]. En definitiva, és plausible que la utilitat d'aquests detectors es limiti a marcar un text que probablement hagi estat generat per IA, contribuint a un conjunt de proves sobre la sospita de frau, en lloc de confirmar la culpabilitat de frau.

2.5 El caràcter disruptiu de la IA

La revisió de les investigacions sobre l'impacte de ChatGPT en l'educació superior, tals com Ojeda et al. (2023) o Lo (2023) posen de manifest el caràcter disruptiu d'aquesta nova tecnologia, la seva capacitat d'evolucionar ràpidament i introdueixen de forma preliminar el desafiament que suposa per a la integritat acadèmica. Desafortunadament, no hem pogut constatar la publicació de gaires articles acadèmics centrats en l'impacte de la IA generativa en la integritat acadèmica. D'aquests darrers, podem destacar-ne tres (d'altres arriben a conclusions similars a la d'aquests tres o bé a cap de significativa). El primer d'ells, publicat tan sols unes setmanes més tard de l'aparició de ChatGPT, és el de Susnjak (2022), qui ja comprova les capacitats de la IA per a generar textos indistingibles dels elaborats per humans i fa propostes per a tractar d'abordar possibles actes deshonestos per part dels estudiants, en particular de metodologies online. En una línia similar, Malinka et al. (2023) comproven la capacitat de la IA per a superar exàmens universitaris i diagnostiquen la necessitat de preparar els estudiants amb models docents alternatius. Finalment, Stella Serrano (2023) aprofundeix en els canvis docents i de model d'avaluació que considera seran necessaris en un futur immediat. Entrarem a continuació a analitzar aquesta literatura referenciada.

[2] Podeu trobar informació detallada sobre aquest assumpte aquí: https://openai.com/index/new-ai-classifier-for-indicating-ai-written-text/

2.6 Contribucions científiques de la relació entre IA generativa i integritat acadèmica

Els articles de Malinka et al. (2023), Stella Serrano (2023) i Susnjak (2022) coincideixen en la preocupació per l'impacte de la intel·ligència artificial (IA), especialment d'eines com ChatGPT, en la integritat acadèmica. Aquest debat és central en l'educació superior, ja que la facilitat amb què aquestes tecnologies poden ser utilitzades per fer trampes posa en dubte la validesa dels resultats acadèmics. Malinka et al. (2023) analitzen com ChatGPT pot completar tasques acadèmiques amb un alt nivell de coherència, fent difícil detectar-ne l'ús en exàmens no vigilats. Tot i que reconeixen que la IA té limitacions en comprensió profunda i originalitat, adverteixen que la seva efectivitat en tasques superficials representa una amenaça real per als processos d'avaluació universitaris. Per la seva banda, Stella Serrano (2023) subratlla la necessitat de redefinir els mètodes d'avaluació per garantir-ne la fiabilitat, destacant que la proliferació d'aquestes eines pot comprometre la validesa dels resultats si no s'apliquen mesures preventives. Proposa avaluacions en temps real i projectes col·laboratius per dificultar l'ús indegut de la IA, així com la formació dels estudiants en l'ús ètic d'aquestes tecnologies. D'altra banda, Susnjak (2022) se centra en l'impacte de ChatGPT en els exàmens en línia, alertant sobre l'augment del frau acadèmic que això pot provocar i la necessitat de dissenyar sistemes d'avaluació més segurs per mantenir la confiança en el sistema educatiu.

Aquests estudis també coincideixen en destacar els reptes i limitacions de la IA en l'educació. Malinka et al. (2023) assenyalen que, tot i generar respostes coherents, ChatGPT no pot substituir el pensament crític i la creativitat necessaris en tasques acadèmiques avançades. Stella Serrano (2023) reforça aquesta idea, argumentant que la IA té dificultats per captar la profunditat dels conceptes acadèmics, la qual cosa en limita el paper com a eina educativa. De manera similar, Susnjak (2022) destaca que, tot i que ChatGPT pot proporcionar respostes correctes en exàmens, li manca capacitat d'anàlisi crítica. En conjunt, aquests autors adverteixen que, encara que la IA pot ser útil per a tasques senzilles, no pot substituir completament els processos educatius tradicionals, i subratllen la necessitat d'integrar-la de manera que complementi l'ensenyament en lloc de reemplaçar-lo.

Un altre punt d'acord entre els articles és la necessitat d'adaptar els mètodes d'avaluació per preservar la integritat acadèmica. Malinka et al. (2023) proposen utilitzar tecnologies de detecció de plagi i redefinir les avaluacions per fomentar l'originalitat i el pensament crític. Stella Serrano (2023) suggereix proves en temps real i sistemes de seguiment d'activitat per evitar l'ús indegut de la IA. Susnjak (2022), per la seva part, aposta per sistemes de vigilància més avançats i per la creació d'avaluacions que dificultin el frau acadèmic. Així, tots tres autors coincideixen en la importància de dissenyar proves que fomentin la participació activa dels estudiants i redueixin la possibilitat de trampes mitjançant la IA.

A més, els tres estudis coincideixen en la importància de promoure un ús ètic de la IA en l'educació. Malinka et al. (2023) defensen la necessitat de formar els estudiants en l'ús responsable d'aquestes tecnologies, incloent-hi cursos sobre ètica digital en el currículum universitari. Stella Serrano (2023) destaca que fomentar una cultura acadèmica basada en la integritat pot ajudar a prevenir el frau acadèmic, mentre que Susnjak (2022) remarca que l'educació ètica ha de complementar les mesures tecnològiques per evitar abusos. D'aquesta manera, tots tres autors insisteixen en la necessitat que les institucions educatives assumeixin un paper actiu en la regulació i educació sobre la IA.

Finalment, aquests estudis proposen reformes per adaptar el sistema educatiu als reptes de la IA. Malinka et al. (2023) recomanen l'ús de metodologies d'ensenyament més interactives i col·laboratives, mentre que Stella Serrano (2023) aposta per actualitzar els currículums i formar el professorat en l'ús de la tecnologia. Susnjak (2022), per la seva banda, defensa l'ús de plataformes d'aprenentatge que incorporin sistemes de vigilància intel·ligents per detectar el frau acadèmic. En conjunt, tots tres autors assenyalen que la resposta als desafiaments de la IA ha de passar per la innovació en els mètodes d'ensenyament i avaluació, així com per la implementació de polítiques clares i efectives per garantir que la IA s'integri de manera beneficiosa en l'educació superior.

3. Conclusions i futures línies de recerca

Tal i com dèiem, la IA planteja la necessitat imperiosa de repensar els sistemes d'avaluació dels aprenentatges i, especialment, el seu caràcter formatiu. És indispensable dur a terme una revisió dels sistemes i objectius de l'avaluació per centrar-los a avaluar prioritàriament processos i no únicament productes. En especial, s'han de replantejar la conveniència i el disseny de les activitats que requereixen més treball autònom i determinar criteris per acreditar competències. A continuació, ampliem aquest apartat enumerant i detallant les conclusions que es desprenen de la nostra recerca i les mesures que es poden aplicar per millorar la integritat acadèmica en l'avaluació de les activitats formatives universitàries.

1. Actualització de Normatives i Codis Ètics: les universitats han de revisar i actualitzar els seus codis ètics i normatives d'integritat acadèmica per incloure les noves formes de frau facilitades per la IA generativa. Aquesta actualització és urgent, ja que alguns dels documents actuals daten d'abans de l'era de la IA generativa i no contemplen els seus usos i implicacions, i en els que són més recents sovint no se'n fa cap menció. El caràcter altament disruptiu d'aquesta tecnologia obliga a evitar ambigüitats i inconcrecions. Tal i com ha mostrat la literatura científica estudiada, les universitats han d'establir i comunicar clarament polítiques de deshonestedat acadèmica, incloent definicions específiques de comportaments deshonestos i les conseqüències de ser descobert. Aquestes polítiques han de ser conegudes tant pels estudiants com pels professors. Una definició clara de la casuística potencialment detectada i els procediments que se'n derivaran - per exemple, que en cas de dubte de frau l'alumnat serà entrevistat pel professorat o fins i tot avaluat oralment - pot actuar com a element dissuasiu en l'ús de pràctiques deshonestes. Per la seva banda, les agències de qualitat i altres organismes públics de supervisió haurien de contribuir activament en l'articulació d'aquestes polítiques, alhora que haurien de validar en els seus processos d'acreditació l'adequació dels mecanismes implementats singularment a cada institució.

2. Canvi en la Metodologia Docent i Avaluativa: cal adoptar mètodes d'avaluació que requereixin una participació activa dels estudiants, com avalu-

acions en temps real (*in class*), projectes col·laboratius, treball de camp, discussió de casos pràctics i preguntes que impliquin aplicació del coneixement i solució de problemes. L'adopció d'aquestes mesures requereixen d'una dedicació i un seguiment més permanent dels docents i no és contrària al propi ús de la IA generativa. De fet, hauria de permetre que aquesta es convertís en una eina més - o la més important - a disposició de l'alumnat, qui hauria de detectar que la IA ajuda, però no soluciona o no proporciona les solucions que busca l'usuari. No es tracta només d'una aproximació *raise the bar*, això és, de pressuposar que tothom està usant la IA generativa i per tant el que es demana com a compliment és notablement superior al que es demanava abans que se'n pogués fer ús, sinó que directament es canvien les preguntes i els ítems avaluables. Conceptes com treballar les *soft skills*, el pensament crític, la creativitat o pensar *out of the box* adquireixen aquí una nova centralitat, relegant els continguts a la fonamentació d'aquestes competències. Aquests canvis en la metodologia són especialment profunds i costosos a nivell de recursos humans - probablement es requereixin més docents i més dedicació -, però dificulten o directament eliminen l'ús de la IA generativa per cometre frau acadèmic.

3. Foment de la Cultura d'Integritat Acadèmica: ja hem assenyalat com les institucions catalanes han prestat atenció especialment a fer pedagogia de bones pràctiques de la IA generativa i aquesta línia cal que continuï i que sigui reforçada. Per exemple, es pot avançar en programes de mentoria on estudiants més grans o tutors ajuden els nous estudiants a entendre les expectatives acadèmiques i a desenvolupar habilitats de gestió del temps i estudi que redueixin la temptació de fer trampes. També hauríem d'incloure aquí les campanyes de conscienciació per promoure un comportament ètic entre els estudiants o les sessions de formació obligatòries sobre integritat acadèmica per a estudiants i professors centrades en la comprensió de què constitueix la deshonestedat acadèmica i com evitar-la. En definitiva, es tracta no només de penalitzar les actituds deshonestes de manera efectiva, sinó de fomentar una cultura institucional que valori la integritat acadèmica, reconeixent i recompensant el comportament ètic, així com creant un ambient on els estudiants se sentin còmodes denunciant casos de deshonestedat acadèmica.

4. Ús Estratègic de les Tecnologies: eines com el campus virtual poden ser útils per avaluar processos i no únicament productes, així com per establir criteris per acreditar competències. L'ús estratègic de les tecnologies educatives pot ajudar a monitoritzar el procés d'aprenentatge de l'alumnat i a registrar-ne els avenços. Fins i tot, aquest enregistrament codificat podria arribar a ser reavaluat mitjançant una futura tecnologia de detecció de plagi o d'actitud deshonesta i actuant com a desincentiu, tal i com ens ensenyava la literatura acadèmica. És cert, no obstant, que l'evolució pròpia de la IA generativa ocasionarà nous reptes que ara només podem intuir i que requeriran d'una capacitat de reacció àgil i eficaç per part de les universitats. Per exemple, no es descartable que avaluacions fetes per videoconferència no puguin detectar l'ús d'avatars indistingibles en aparença de l'alumne que s'està examinant i que donen respostes generades immediatament per IA generativa. De la mateixa manera, agents IA degudament entrenats podrien simular de forma versemblant un avenç progressiu en el procés d'aprenentatge dins el campus virtual.

5. Avaluació Continua: En lloc de dependre únicament d'avaluacions sumatòries al final del curs, hauria de ser recomanable promoure l'avaluació contínua, implementant avaluacions formatives durant tot el curs. Això pot reduir la pressió sobre els estudiants i proporcionar oportunitats per al *feedback* i la millora. No obstant, l'avaluació contínua - i més si s'usa una varietat de mètodes d'avaluació, com ara projectes en grup, presentacions orals, diaris reflexius i exàmens pràctics - agreuja una de les preocupacions tradicionals de les universitats, que és l'accentuació de la subjectivitat i que pot desembocar en sinònim de conflictivitat. Les universitats acostumen a fer esforços per establir unes rúbriques d'avaluació i uns criteris pretesament objectius que determinen una qualificació final. L'avaluació continua - i els canvis metodològics proposats en el punt 2 d'aquest epígraf - dificulta especialment l'objectivitat d'allò que s'està avaluant, en especial quan no es tracta d'un entregable registrat, sinó d'una discussió en un debat, una avaluació oral o un procés d'aprenentatge. És per aquest motiu que els exàmens orals acostumen a fer-se davant més d'un professor - generalment un número senar d'avaluadors -. Caldrà, doncs, tenir en compte aquestes consideracions si es volen implementar canvis substancials en els processos d'avaluació acadèmica.

6. Supervisió Estricta: en els exàmens en línia, s'han d'utilitzar programes de supervisió i seguiment que monitoritzin l'activitat de l'ordinador dels estudiants durant l'examen per detectar comportaments sospitosos. Aquesta mesura és essencial per garantir la integritat dels exàmens en línia i per dissuadir els estudiants de cometre frau acadèmic. Malgrat tot, és probable que alguns alumnes trobin la manera de sortejar aquesta supervisió. A més, mesures pròpies de l'entorn en línia com dissenyar exàmens amb preguntes de temps limitat i que demanin una anàlisi i una aplicació - per exemple, la resolució d'un cas pràctic - poden superar-se fàcilment amb la intervenció d'una IA generativa. Per aquest motiu, és recomanable aprofitar alguns dels canvis metodològics proposats i fer referència en els exàmens a qüestions plantejades dins l'aula o en sessions síncrones i que siguin prou específiques i de naturalesa crítica per evitar l'ús fraudulent de la IA generativa.

En definitiva, les universitats i les comunitats educatives en general han de prendre mesures proactives per adaptar-se als canvis que la IA generativa introdueix en l'educació superior. Casos recents com el del Tribunal Superior de Justícia rus, que al març del 2023 va considerar original un Treball fi d'estudis elaborat íntegrament amb IA generativa avisen sobre la complexitat del moment, amb ramificacions que escapen l'abast d'aquesta recerca. En qualsevol cas, a través d'una combinació de les estratègies esmentades, es pot fomentar una cultura d'integritat acadèmica i reduir la incidència de comportaments deshonestos en l'avaluació de les activitats formatives universitàries. La implementació d'aquestes mesures és crucial per garantir la qualitat i la validesa dels aprenentatges avaluats en un context on la IA generativa està transformant els processos educatius.

Referencies citades

Abd-Elaal, E.-S., Gamage, S. H. P. W., & Mills, J. E. (2019). Artificial intelligence is a tool for cheating academic integrity. *Proceedings of the 30th Annual Conference for the Australasian Association for Engineering Education (AAEE 2019): Educators Becoming Agents of Change: Innovate, Integrate, Motivate*, Brisbane, Australia, 8-11 December 2019.

Akbulut, Y., Şendağ, S., Birinci, G., Kılıçer, K., Şahin, M. C., & Odabaşı, H. F. (2008). Exploring the types and reasons of Internet-triggered academic dishonesty among Turkish undergraduate

students: Development of Internet-Triggered Academic Dishonesty Scale (ITADS). *Computers & Education, 51*(1), 463-473. https://doi.org/10.1016/j.compedu.2007.06.003

Awdry, R., & Ives, B. (2021). Students cheat more often from those known to them: Situation matters more than the individual. *Assessment & Evaluation in Higher Education, 46*(8), 1254-1268. https://doi.org/10.1080/02602938.2020.1851651

Becker, D., Connolly, J., Lentz, P., & Morrison, J. (2006). Using the Business Fraud Triangle to Predict Academic Dishonesty among Business Students. *The Academy of Educational Leadership Journal.* https://www.semanticscholar.org/paper/Using-the-Business-Fraud-Triangle-to-Predict-among-Becker-Connolly/b6fd4e20a10b3577cd0abc2e04a5b6d71054efb8

Birks, D., & Clare, J. (2023). Linking artificial intelligence facilitated academic misconduct to existing prevention frameworks. *International Journal for Educational Integrity, 19*(1), 20. https://doi.org/10.1007/s40979-023-00142-3

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., Rozenberg, P., Saddiqui, S., & van Haeringen, K. (2019). Contract cheating: A survey of Australian university students. *Studies in Higher Education, 44*(11), 1837-1856. https://doi.org/10.1080/03075079.2018.1462788

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT (arXiv:2303.04226). arXiv. http://arxiv.org/abs/2303.04226

Curtis, G., & Vardenaga, R. (2016). Is plagiarism changing over time? A 10-year time-lag study with three points of measurement. *Higher Education Research & Development, 35*(6). https://www.tandfonline.com/doi/full/10.1080/07294360.2016.1161602

García-Peña, V. R., Mora-Marcillo, A. B., & Ávila-Ramírez, J. A. (2020). La inteligencia artificial en la educación. *Apuntes Universitarios, 12*(1).

Hinman, L. (2002). Academic integrity and the World Wide Web. *ACM SIGCAS Computers and Society, 32*, 33-42. https://doi.org/10.1145/511134.511139

Hodgkinson, T., Curtis, H., MacAlister, D., & Farrell, G. (2016). Student Academic Dishonesty: The Potential for Situational Prevention. *Journal of Criminal Justice Education, 27*(1), 1-18. https://doi.org/10.1080/10511253.2015.1064982

Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic Integrity in Online Assessment: A Research Review. *Frontiers in Education, 6*, 639814. https://doi.org/10.3389/feduc.2021.639814

Incio Flores, F. A., Capuñay Sanchez, D. L., Estela Urbina, R. O., Valles Coral, M. Á., Vergara Medrano, E. E., & Elera Gonzales, D. G. (2021). Inteligencia artificial en educación: Una revisión de la literatura en revistas científicas internacionales. *Apuntes Universitarios, 12*(1). https://doi.org/10.17162/au.v12i1.974

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns, 4*(7), 100779. https://doi.org/10.1016/j.patter.2023.100779

Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the

Literature. *Education Sciences*, 13(4), 410. https://doi.org/10.3390/educsci13040410

Malinka, K., Peresíni, M., Firc, A., Hujnák, O., & Janus, F. (2023). On the Educational Impact of ChatGPT: Is Artificial Intelligence Ready to Obtain a University Degree? *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, 47-53. https://doi.org/10.1145/3587102.3588827

Mut, B. (2010). Approach to avoiding plagiarism strategies developed by Universities of East Anglia (Norwich) and John Moore's University (Liverpool).

Ojeda, A. D., Solano-Barliza, A. D., Alvarez, D. O., & Cárcamo, E. B. (2023). Análisis del impacto de la inteligencia artificial ChatGPT en los procesos de enseñanza y aprendizaje en la educación universitaria. *Formación universitaria, 16*(6), 61-70. https://doi.org/10.4067/S0718-50062023000600061

Pedreño, A., González, R., Mora, T., Pérez, E. M., Ruiz, J., & Torres, A. (2024). *La inteligencia artificial en las universidades: retos y oportunidades.* Grupo 1MillionBot.

Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2024). Can AI-Generated Text be Reliably Detected? (*arXiv:2303.11156*). arXiv. http://arxiv.org/abs/2303.11156

Stella Serrano, C. (2023). La llegada de la inteligencia artificial y el problema de la evaluación en la docencia universitaria. *El sistema educativo en crisis.* https://repositorio.uam.es/handle/10486/711916

Susnjak, T. (2022). ChatGPT: The End of Online Exam Integrity? (arXiv:2212.09292). *arXiv.* http://arxiv.org/abs/2212.09292

Swauger, S. (2020). Software that monitors students during tests perpetuates inequality and violates their privacy. *MIT Technology Review.* https://www.technologyreview.com/2020/08/07/1006132/software-algorithms-proctoring-online-tests-ai-ethics/

Tolman, S. (2017). Academic Dishonesty in Online Courses: Considerations for Graduate Preparatory Programs in Higher Education. *College Student Journal, 51*(4), 579-584.

Wahid, A., Sengoku, Y., & Mambo, M. (2015). Toward constructing a secure online examination system. *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*, 1-8. https://doi.org/10.1145/2701126.2701203

Weber-Wulff, D., et al. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19*(1), 26. https://doi.org/10.1007/s40979-023-00146-z